N080-630

ERIC REPORT RESUME

ED 010 044 9-01-66 24 (REV)
THE CONSTRUCTION AND VALIDATION OF TESTS OF THE COGNITIVE PROCESSES AS DESCRIBED IN THE 'TAXONOMY OF EDUCATIONAL OBJECTIVES'.
KROPP, RUSSELL P. * AND OTHERS
EJI15118 FLORIDA STATE UNIV., INST. OF HUMAN LEARNING, TALLAHASSEE
CRP-2117
-FEB-66 OEC-4-10-019
EDRS PRICE MF-$0.63 HC-$17.76 444P.

*TAXONOMY, *VALIDATION, *EVALUATION, *RESPONSE DEVICES, *COGNITIVE DEVELOPMENT, TESTING, TALLAHASSEE, FLORIDA

THE BROAD PURPOSE OF THIS PROJECT WAS TO TEST EMPIRICALLY THE STRUCTURE OF THE HANDBOOK 'TAXONOMY OF EDUCATIONAL OBJECTIVES 1--COGNITIVE DOMAIN.' IT WAS NECESSARY TO CONSTRUCT FOUR FORMS OF A TAXONOMY-TYPE TEST SINCE NONE WAS AVAILABLE. THE PHENOMENA WHICH ARE CLASSIFIED BY THE 'TAXONOMY' ARE THE RESPONSES (BEHAVIORS) WHICH TEST ITEMS (IN THE COGNITIVE DOMAIN AS CONTRASTED WITH THE AFFECTIVE AND PSYCHOMOTOR DOMAINS) ARE INTENDED TO EVOKE. AS AN AID TO THE READER, THE REPORT CONTAINS SEVERAL SECTIONS WHICH DEAL ENTIRELY WITH THE TEST DEVELOPMENT AND TEST ADMINISTRATION. THE VALIDITY OF THE TAXONOMIC STRUCTURE OF IMPUTED HIERARCHY WAS GENERALLY SUPPORTED BY THE EMPIRICAL DATA. FURTHER STUDY WAS RECOMMENDED IN SUCH AREAS AS KNOWLEDGE AS A PROCESS, ITEM ANALYSIS DATA, TRIANGULAR BIVARIATE DISTRIBUTIONS, INTENDED AND ACTUAL PROCESSES, THE EVALUATION PROCESS, NORMS FOR TAXONOMY-TYPE TESTS, ITEM TYPES, AND PYRAMID TESTS. (GD)

THE CONSTRUCTION AND VALIDATION OF TESTS OF THE COGNITIVE PROCESSES AS DESCRIBED IN THE TAXONOMY OF EDUCATIONAL OBJECTIVES

Cooperative Research Project No. 2117

Russell P. Kropp
Howard W. Stoker

with the assistance of
W. Louis Bashaw

Institute of Human Learning
and
Department of Educational Research and Testing
Florida State University

February, 1966

The research reported herein was supported by the Cooperative Research Program of the Office of Education, U.S. Department of Health, Education, and Welfare.

I

PRECEDING PAGE BLANK-NOT FILMED

PREFACE

This report is an account of a three-year long series of studies to explore the construct validity of the Taxonomy of Educational Objectives, Handbook I; Cognitive Domain.

A special note about the structure of this report should be made at the outset. The broad purpose of the study was to test empirically the structure of the taxonomy. To do this it was necessary to construct four forms of a taxonomy-type test because none was otherwise available. The report contains several sections which deal entirely with the test development and test administration. Ordinarily, a report does not so heavily emphasize instrument development unless it is in itself the research goal. It is not in this case but we did so for several reasons. We regard the taxonomy as a potentially valuable research device and as a comprehensive description of a kind of educational objective which is critically important but quite ignored when educational progress is evaluated. Thus, we believe that presenting the tests and an account of their development might lead others to use them experimentally and to develop vastly better versions of them. Also, we identified some fundamental problems during test development which are most clearly presented when embedded in the context in which they were found. The identification of these problems and others which arose concerning the legitimate uses of data from the tests might jointly constitute whatever contribution, if any, which inheres in this report. Our goal was the validation of the Taxonomy but our accomplishment might be nothing more than having laid bare problems which must be surmounted before that goal can seriously be pursued.

The preliminary studies were supported by the Department of Educational Research and Testing of the Florida State University and by the Research Council of the University. The latter was in the form of grant-in-aid to Dr. R. P. Kropp and Dr. H. W. Stoker for "Exploratory Studies in the Validation and Development of Measures of Cognitive Processes." The major study was supported by the U. S. Office of Education, Cooperative Research Program, and the Institute of Human Learning and the Department of Educational Research and Testing of the Florida State University.

The Department and Institute encourage widespread faculty and student participation in research projects. The following categories of people participated in the overall project at one time or another during its duration.

Project faculty. Dr. Russell P. Kropp and Dr. Howard Stoker served as co-principal investigators. Dr. W. Louis Bashaw, Assistant Professor of Educational Research and Testing, served as a member of the project faculty from its inception until December 31, 1965, at which time he joined the faculty of the University of Georgia.

Participating faculty. Dr. Thomas K. Bullock, Department of Foundations of Education, and Mr. Harry D. Jackson, Department of Social Studies Education, assisted in selecting reading passages for the taxonomy-type tests which were constructed. Dr. Ernestine O'Connell, who was then a member of the faculty of the Department of Science Education and is now on the faculty of Old Dominion College, Norfolk, Virginia, assisted in selecting and abridging reading passages and in writing, editing, and trying-out items. Drs. Donald L. Hartford, Edward L. Palmer, and Ian E. Reid, all assistant professors in the Department of Educational Research and Testing, advised on the general methodology of the study and the statistical analyses of the resulting data. Dr. F. J. King, Associate Professor of Educational Research and Testing and Research Associate in the Institute of Human Learning, consulted on and assisted with the analyses of data.

Project graduate assistants. Mr. Howard R. Inglis, now of Ohio State University, assisted primarily in test construction and test administration. Dr. Emmett T. Kohler, now a faculty member at the University of Georgia, assisted primarily in test construction and data analyses. Mrs. Katherine A. McClure assisted in test administration. Mr. Lawrence Moore, now an Education and Training Specialist with General Electric in Daytona Beach, Florida, assisted primarily in test construction and administration. Mr. Isaiah Moyel, now a faculty member at Columbus College, Georgia, assisted in test construction. Mr. C. Neil Shaw assisted in test construction, test administration, data analyses, and the preparation of scoring procedures for Evaluation and Synthesis items. Mr. Richard Sowash conducted oral problem solving studies to determine item validities. Mr. William R. Turner assisted in test construction, test administration, and data analyses. Mr. John L. Wasik, now Director of Research for the Montgomery County Schools, Maryland, assisted in test construction, test administration, and data analyses.

Non-project graduate students. Dr. June S. Anderson, now a faculty member at Middle Tennessee State College, Murfreesboro, Tennessee, assisted in the development of taxonomy-type tests in the sciences. Dr. Peter Dunn-Rankin, now a faculty member at the University of Hawaii, and Mr. James R. Swanson assisted in test construction. Mrs. Coleen Story assisted in test construction.

Consultants. Dr. J. Thomas Hastings, Director, Center for Instructional Research and Curriculum Evaluation, University of Illinois, and Dean David R. Krathwohl, Syracuse University, served as general consultants.

Although he was not a consultant to the project staff, we wish to acknowledge the perceptive comments and advice which were given to us by Dr. Joe Saupe of Michigan State University.

The project required extensive data collection from several thousand public school students. This phase was admirably facilitated by the superintendents of public instruction and the directors of guidance, testing, and research in each of the following counties, all of which are in Florida: Brevard (Titusville), Broward (Ft. Lauderdale), Duval (Jacksonville), Hillsborough (Tampa), Madison (Madison), Manatee (Bradenton), Polk (Bartow), and Volusia (DeLand). During test administrations, teachers from secondary schools in these counties served as proctors and in that role invaluably contributed to the project. We wish to acknowledge with genuine gratitude the help which all these people so willingly gave.

We wish to acknowledge our appreciation to the Educational Testing Service for permitting us to reproduce tests which appear in the Kit of Reference Tests.

We wish to express our gratitude to the several authors and publishing companies who permitted us to adapt their copyrighted materials for use in the reading passages which are part of the taxonomy-type tests we constructed. Specific acknowledgments of these permissions appear in Appendices A through D which contain the experimental test forms.

We wish also to express our appreciation to the Florida State University Computing Center staff members for their many kind services and to the National Science Foundation for its support of the Computing Center.

Russell P. Kropp
Howard W. Stoker
February 28, 1966

PRECEDING PAGE BLANK-NOT FILMED

TABLE OF CONTENTS

LIST OF TABLES

Number		Page
62a	Summary of Predictors of Level Scores, over Forms by Grades with Significant Beta Weights, Grade 11	144
63	Summary of Significant Beta Weights, Grade 12	145
63a	Summary of Predictors of Level Scores, over Forms by Grades with Significant Beta Weights, Grade 12	146
64	Named KIT Factors by Grade	147

I. PROBLEM

This section consists of a description of the Taxonomy of Educational Objectives, Handbook I: Cognitive Domain (Bloom, 1956), a statement of the problem which was studied, and some remarks about its significance.

The description of the Taxonomy is somewhat extended because it constitutes the single document of interest in the study. Particular attention is devoted to those parts of it which deal with the rationale for the structure of the taxonomy and which deal with the taxonomists' conception of cognitive aptitude or ability and its generality. The condensed version of the Taxonomy appears in this section as an aid to the reader and through the gracious permission of the publisher.

A. Taxonomy of Educational Objectives, Handbook I: Cognitive Domain

Part I of the Taxonomy describes the history of the development of the taxonomy. The reader who desires a comprehensive account should refer to it. What follows is a condensation of those parts of it which are most relevant to this study.

In 1948 a group of college examiners conceived the idea of developing a classification system for educational objectives. The purpose of it would be to facilitate communication among them about test items, educational objectives, and testing procedures. Hitherto, communication among them was impaired by the lack of a standardized vocabulary and conceptual framework.

The group conceived three classification systems, each in a different domain of behavior; cognitive, affective, and psychomotor. The first to be completed was the cognitive domain and one thousand copies of it were printed and distributed in a preliminary edition. It was distributed to public school and college educators, and others with the request that it be criticized. These criticisms were then taken into account in preparing for publication the version which is cited here.

Several decisions which the college examiners made in order to prepare the taxonomy are relevant to this study. These are described briefly below but more complete discussions of them appear in subsequent parts of the report.

The developers agreed that as a first priority the classification scheme or taxonomy should reflect distinctions teachers make about student behaviors. As a second priority, they agreed that the taxonomy should be logically developed and internally consistent. Third, the classification scheme should be consistent with present understanding of psychological phenomena. And, finally, the classification scheme should be descriptive and it should be neutral as regards educational philosophies. Some implications of these discussions are mentioned in the Taxonomy (Bloom, 1956, pp. 13-15). It is noteworthy that prime attention was accorded producing a taxonomy which would be educationally meaningful. This emphasis is congruent with the difficulties they had previously experienced in communicating about test problems. Since this was a conscious decision, and seemingly a very sensible and defensible one, the developers were certainly aware that in some instances logical and psychological relevance, technical excellence, and comprehensiveness might have to be partially sacrificed to achieve educational meaningfulness.

They had to choose the phenomena on which to develop the taxonomy. Their choice was the student behavior which a test item is intended to elicit. In reaching this decision, they discarded other phenomena like subject matter, educational processes, and, more especially, student behavior which an item actually evokes. In this latter regard, they acknowledged that the behaviors which an item is intended to evoke and those which it actually evokes might be different due to prior experiences of the examinees.

The taxonomists then decided that the intended behaviors or processes could be categorized into broad groups and within each group into specific subgroups. They then arranged these broad groups of behaviors or processes in what they believed to be an hierarchical structure.

The structure was presumed to be hierarchical on the basis that behaviors which defined the first broad category were regarded as being integrated with new behaviors in the next higher category, and that the integration of these behaviors and new behaviors represented the next higher category, etc. In addition, the specific subgroups of a broad group were arranged according to complexity, if not hierarchically.

The resulting taxonomy, in a condensed form, appears below with the permission of the publisher and is taken intact from the Taxonomy (Bloom, 1956, pp. 201-207).

Condensed Version of Taxonomy of Educational Objectives

Cognitive Domain

KNOWLEDGE

1.00 KNOWLEDGE

Knowledge, as defined here, involves the recall of specifics and universals, the recall of methods and processes, or the recall of a pattern, structure, or setting. For measurement purposes, the recall situation involves little more than bringing to mind the appropriate material. Although some alteration of the material may be required, this is a relatively minor part of the task. The knowledge objectives emphasize most the psychological processes of remembering. The process of relating is also involved in that a knowledge test situation requires the organization and reorganization of a problem such that it will furnish the appropriate signals and cues for the information and knowledge the individual possesses. To use an analogy, if one thinks of the mind as a file, the problem in a knowledge test situation is that of finding in the problem or task the appropriate signals, cues, and clues which will most effectively bring out whatever knowledge is filed or stored.

1.10 KNOWLEDGE OF SPECIFICS

The recall of specific and isolable bits of information. The emphasis is on symbols with concrete referents. This material, which is at a very low level of abstraction, may be thought of as the elements from which more complex and abstract forms of knowledge are built.

1.11 KNOWLEDGE OF TERMINOLOGY

Knowledge of the referents for specific symbols (verbal and non-verbal). This may include knowledge of the most generally accepted symbol referent, knowledge of the variety of symbols which may be used for a single referent, or knowledge of the referent most appropriate to a given use of a symbol.

*To define technical terms by giving their attributes, properties, or relations.

*Familiarity with a large number of words in their common range of meanings.

1.12 KNOWLEDGE OF SPECIFIC FACTS

Knowledge of the dates, events, persons, places, etc. This may include very precise and specific information such as the specific date or exact magnitude of a phenomenon. It may also include approximate or relative information such as an approximate time period or the general order of magnitude of a phenomenon.

*The recall of major facts about particular cultures.

*The possession of a minimum knowledge about the organisms studied in the laboratory.

1.20 KNOWLEDGE OF WAYS AND MEANS OF DEALING WITH SPECIFICS

Knowledge of the ways of organizing, studying, judging, and criticizing. This includes the methods of inquiry, the chronological sequences, and the standards of judgment within a field as well as the patterns of organization through which the areas of the fields themselves are determined and internally organized. This knowledge is at an intermediate level of abstraction between specific knowledge on the one hand and knowledge of universals on the other. It does not so much demand the activity of the student in using the materials as it does a more passive awareness of their nature.

1.21 KNOWLEDGE OF CONVENTIONS

Knowledge of characteristic ways of treating and presenting ideas and phenomena. For purposes of communication and consistency, workers in a field employ usages, styles, practices, and forms which best suit their purposes and/or which appear to

---

*Illustrative educational objectives selected from the literature.

suit best the phenomena with which they deal. It should be recognized that although these forms and conventions are likely to be set up on arbitrary, accidental, or authoritative bases, they are retained because of the general agreement or concurrence of individuals concerned with the subject, phenomena, or problem.

*Familiarity with the forms and conventions of the major types of works, e.g., verse, plays, scientific papers, etc.

*To make pupils conscious of correct form and usage in speech and writing.

1.22 KNOWLEDGE OF TRENDS AND SEQUENCES

Knowledge of the processes, directions, and movements of phenomena with respect to time.

*Understanding of the continuity and development of American culture as exemplified in American life.

*Knowledge of the basic trends underlying the development of public assistance programs.

1.23 KNOWLEDGE OF CLASSIFICATIONS AND CATEGORIES

Knowledge of the classes, sets, divisions, and arrangements which are regarded as fundamental for a given subject field, purpose, argument, or problem.

*To recognize the area encompassed by various kinds of problems or materials.

*Becoming familiar with a range of types of literature.

1.24 KNOWLEDGE OF CRITERIA

Knowledge of the criteria by which facts, principles, opinions, and conduct are tested or judged.

*Familiarity with criteria for judgment appropriate to the type of work and the purpose for which it is read.

*Knowledge of criteria for the evaluation of recreational activities.

1.25 KNOWLEDGE OF METHODOLOGY

Knowledge of the methods of inquiry, techniques, and procedures employed in a particular subject field as well as those employed in investigating particular problems and phenomena. The emphasis here is on the individual's knowledge of the method rather than his ability to use the method.

*Knowledge of scientific methods for evaluating health concepts.

*The student shall know the methods of attack relevant to the kinds of problems of concern to the social sciences.

1.30 KNOWLEDGE OF THE UNIVERSALS AND ABSTRACTIONS IN A FIELD

Knowledge of the major schemes and patterns by which phenomena and ideas are organized. These are the large structures, theories, and generalizations which dominate a subject field or which are quite generally used in studying phenomena or solving problems. These are at the highest levels of abstraction and complexity.

1.31 KNOWLEDGE OF PRINCIPLES AND GENERALIZATIONS

Knowledge of particular abstractions which summarize observations of phenomena. These are the abstractions which are of value in explaining, describing, predicting, or in determining the most appropriate and relevant action or direction to be taken.

*Knowledge of the important principles by which our experience with biological phenomena is summarized.

*The recall of major generalizations about particular cultures.

1.32 KNOWLEDGE OF THEORIES AND STRUCTURES

Knowledge of the body of principles and generalizations together with their interrelations which present a clear, rounded, and systematic view of a complex phenomenon, problem or field. These are the most abstract formulations, and they can be used to show the interrelation and organization of a great range of specifics.

*The recall of major theories about particular cultures.

*Knowledge of a relatively complete formulation of the theory of evolution.

INTELLECTUAL ABILITIES AND SKILLS

Abilities and skills refer to organized modes of operation and generalized techniques for dealing with materials and problems. The materials and problems may be of such a nature that little or no specialized and technical information is required. Such information as is required can be assumed to be part of the individual's general fund of knowledge. Other problems may require specialized and technical information at a rather high level such that specific knowledge and skill in dealing with the problem and the materials are required. The abilities and skills objectives emphasize the mental processes of organizing and reorganizing material to achieve a particular puppose. The materials may be given or remembered.

2.00 <u>COMPREHENSION</u>

This represents the lowest level of understanding. It refers to a type of understanding or apprehension such that the individual knows what is being communicated and can make use of the material or idea being communicated without necessarily relating it to other material or seeing its fullest implications.

2.10 <u>TRANSLATION</u>

Comprehension as evidenced by the care and accuracy with which the communication is paraphrased or rendered from one language or form of communication to another. Translation is judged on the basis of faithfulness and accuracy, that is, on the extent to which the material in the original communication is preserved although the form of the communication has been altered.

*The ability to understand non-literal statements (metaphor, symbolism, irony, exaggeration).

*Skill in translating mathematical verbal material into symbolic statements and vice versa.

2.20 INTERPRETATION

The explanation or summarization of a communication. Whereas translation involves an objective part-for-part rendering of a communication, interpretation involves a reordering, rearrangement, or a new view of the material.

*The ability to grasp the thought of the work as a whole at any desired level of generality.

*The ability to interpret various types of social data.

2.30 EXTRAPOLATION

The extension of trends or tendencies beyond the given data to determine implications, consequences, corollaries, effects, etc., which are in accordance with the conditions described in the original communication.

*The ability to deal with the conclusions of a work in terms of the immediate inference made from the explicit statements.

*Skill in predicting continuation of trends.

3.00 APPLICATION

The use of abstractions in particular and concrete situations. The abstractions may be in the form of general ideas, rules of procedures, or generalized methods. The abstractions may also be technical principles, ideas, and theories which must be remembered and applied.

*Application to the phenomena discussed in one paper of the scientific terms or concepts used in other papers.

*The ability to predict the probable effect of a change in a factor on a biological situation previously at equilibrium.

4.00 ANALYSIS

The breakdown of a communication into its constituent elements or parts such that the relative hierarchy of ideas is made clear and/or the relations between

the ideas expressed are made explicit. Such analyses are intended to clarify the communication, to indicate how the communication is organized, and the way in which it manages to convey its effects, as well as its basis and arrangement.

4.10 ANALYSIS OF ELEMENTS

Identification of the elements included in a communication.

*The ability to recognize unstated assumptions.

*Skill in distinguishing facts from hypotheses.

4.20 ANALYSES OF RELATIONSHIPS

The connections and interactions between elements and parts of a communication.

*Ability to check the consistency of hypotheses with given information and assumptions.

*Skill in comprehending the interrelationships among the ideas in a passage.

4.30 ANALYSIS OF ORGANIZATIONAL PRINCIPLES

The organization, systematic arrangement, and structure which hold the communication together. This includes the "explicit" as well as "implicit" structure. It includes the bases, necessary arrangement, and the mechanics which make the communication a unit.

*The ability to recognize form and pattern in literary or artistic works as a means of understanding their meaning.

*Ability to recognize the general techniques used in persuasive materials, such as advertising, propaganda, etc.

5.00 SYNTHESIS

The putting together of elements and parts so as to form a whole. This involves the process of working with pieces, parts, elements, etc., and arranging and combining them in such a way as to constitute a pattern or structure not clearly there before.

5.10 PRODUCTION OF A UNIQUE COMMUNICATION

The development of a communication in which the writer or speaker attempts to convey ideas, feelings, and/or experiences to others.

*Skill in writing, using an excellent organization of ideas and statements.

*Ability to tell a personal experience effectively.

5.20 PRODUCTION OF A PLAN, OR PROPOSED SET OF OPERATIONS

The development of a plan of work or the proposal of a plan of operations. The plan should satisfy requirements of the task which may be given to the student or which he may develop for himself.

*Ability to propose ways of testing hypotheses.

*Ability to plan a unit of instruction for a particular teaching situation.

5.30 DERIVATION OF A SET OF ABSTRACT RELATIONS

The development of a set of abstract relations either to classify or explain particular data or phenomena, or the deduction of propositions and relations from a set of basic propositions or symbolic representations.

*Ability to formulate appropriate hypotheses based upon an analysis of factors involved, and to modify such hypotheses in the light of new factors and considerations.

*Ability to make mathematical discoveries and generalizations.

6.00 EVALUATION

Judgments about the value of material and methods for given purposes. Quantitative and qualitative judgments about the extent to which material and methods satisfy criteria. Use of a standard of appraisal. The criteria may be those determined by the student or those which are given to him.

6.10 JUDGMENTS IN TERMS OF INTERNAL EVIDENCE

Evaluation of the accuracy of a communication from such evidence as logical accuracy, consistency, and other internal criteria.

*Judging by internal standards, the ability to assess general probability of accuracy in reporting facts from the care given to exactness of statement, documentation, proof, etc.

*The ability to indicate logical fallacies in arguments.

6.20 JUDGMENTS IN TERMS OF EXTERNAL CRITERIA

Evaluation of material with reference to selected or remembered criteria.

*The comparison of major theories, generalizations, and facts about particular cultures.

*Judging by external standards, the ability to compare a work with the highest known standards in its field--especially with other works of recognized excellence.

B. Specific Problems

The study focused on three specific problems: (a) to test the hierarchical structure of the taxonomy; (b) to determine whether the six major processes, aptitudes, or abilities which are described in the taxonomy transcend subject matter content; and (c) to determine the psychological structure of each of these major processes or abilities. Each of these problems is described below.

Hierarchical Structure

The developers of the taxonomy regard its structure as hierarchical. Consider the following.

> We proceeded to divide the cognitive objectives into subdivisions from the simplest behavior to the most complex.
>
> ...so long as the simpler behaviors may be viewed as components of the more complex behaviors,

we can view the educational process as one of building on the simpler bel vior.

In order to find a single place for each type of behavior, the taxonomy must be organized from simple to complex classes of behavior.

Our attempt to arrange educational behaviors from simple to complex was based on the idea that a particular simple behavior may become integrated with other equally simple behaviors to form a more complex behavior. Thus our classifications may be said to be in the form where behaviors of type A form one class, behaviors of type AB form another class, while behaviors of type ABC form still another class. If this is the real order from simple to complex, it should be related to an order of difficulty such that problems requiring behavior A alone should be answered correctly more frequently than problems requiring AB. We have studied a large number of problems occuring in our comprehensive examinations and have found some evidence to support this hypothesis. Thus, problems requiring knowledge of specific facts are generally answered correctly more frequently than problems requiring a knowledge of the universals and abstractions in a field. Problems requiring knowledge of principles and concepts are correctly answered more frequently than problems requiring both knowledge of the principle and some ability to apply it in new situations. ...Our evidence on this is not entirely satisfactory, but there is an unmistakable trend pointing toward a hierarchy of classes of behavior which is in accordance with our present tentative classification of these behaviors (Bloom, 1956, pp. 15-19).

These quotations document the belief of the developers that the taxonomy is cumulatively hierarchical. Consequently this feature is a salient characteristic of the taxonomy and is the point of entry for one who wishes to determine its construct validity. If such a structure cannot be demonstrated, at least relatively, then the usefulness of the taxonomy is clouded.

This claim or intention of a cumulative hierarchy can be investigated providing that data are available from tests which are constructed according to the taxonomy or can be analyzed in accordance with it such that separate scores for each of the six major levels can be obtained. Because tests

of these kinds were not commercially available at the inception of the project, and for that matter are still not available, the initial major undertaking was to construct such instruments. The presumed structure can be used to generate hypotheses about the probable nature of data from such tests. These hypotheses are stated generally here but detailed discussion of them and the adequacy of the statistical tests of them appears later in the report.

First, if a test is constructed so that it contains equal numbers of items corresponding to each of the six major levels of the taxonomy and if the content of all items represent a sample from a described universe of content, then means of level scores should decrease as the complexity of the level increases, i.e., the mean for Knowledge scores should be greater than the mean for Comprehension scores, etc., with Evaluation scores having the smallest mean.

There are certain assumptions on which this hypothesis is based. For the hypotheses to hold, then item complexity must be synonymous with item difficulty, the obtained raw score distributions must closely resemble the true score distributions, and all behaviors which students perform when responding to items must be in the universe of behaviors defined by the taxonomy.

Second, within the restrictions on a test as described in the first hypothesis, an intercorrelation matrix of major level scores should exhibit simplical structure when judged in terms of the claimed structure. This is a reasonable expectation because the developers essentially claim that the structure is as follows:

Category	Abilities
Knowledge	a
Comprehension	a+b
Application	a+b+c
Analysis	a+b+c+d
Synthesis	a+b+c+d+e
Evaluation	a+b+c+d+e+f

Consequently, adjacent categories of the taxonomy should be more highly correlated than non-adjacent pairs of categories and the correlations should decrease as the distance between categories increases.

The technical description of this is as follows, and is taken from Guttman's discussion of the Radex theory (Guttman, 1954). If a perfect simplex exists, then certain

partial correlations will vanish. In particular, $r_{ik.j} = 0$ for $i < j < k$. These partials vanish only if the numerators equal zero, i.e.

$$r_{ik} = r_{ij}r_{jk}, \text{ where } i < j < k$$

In a perfect simplex this relationship holds for all i, j, and k. If this relationship is satisfied, the correlation matrix will be such that the largest values will lie along the upper-left, lower-right diagonal, followed by the next larger values in the adjacent diagonals and the smallest values in the upper-right and lower-left corners of the matrix. The relationship is such that the column totals in a perfect simplex will be smallest at the left and right extremes and highest for the central columns. When the correlation matrix approaches the simplex but does not satisfy the rigorous requirement of the perfect simplex, the matrix may be classed as quasi-simplex.

Transcendence of Processes

When constructing the taxonomy the developers were confronted with the problem of designating the phenomena to be classified. The choice which they made is described in the following paragraph.

> We are of the opinion that although the objectives and test materials and techniques may be specified in an almost unlimited number of ways, the student behaviors involved in these objectives can be represented by a relatively small number of classes. Therefore, this taxonomy is designed to be a classification of the student behaviors which represent the intended outcomes of the educational process. It is assumed that essentially the same class of behavior may be observed in the usual range of subject-matter content, at different levels of education (elementary, high-school, college), and in different schools. Thus, a single set of classifications should be applicable in all these instances (Bloom, 1956, p. 12).

The assertion in the last sentence implies that the abilities, behaviors, or processes described in the taxonomy transcend content. When reduced to the context of a single test item and with great simplication the argument is that the student response to a test item is the interaction of two relatively independent dimensions of behavior, content and process. The student when confronted with a test item must first identify the appropriate content or subject matter which the item problem requires, and then, second, he must use the

relevant process to manipulate the content in order to produce a response to the item. Thus, the assertion is that mental or cognitive operators are applicable to or general over all contents. Analysis in history, Analysis in geometry, and Analysis in foreign language are different so far as the content goes but the cognitive processes which are applied to the content are identical.

An appropriate hypothesis for testing this assertion is that if several taxonomy-type tests are available, each dealing with a different content, and each yielding separate scores for each level of the taxonomy, then the intercorrelations of scores dealing with the same process over several contents will be greater than the correlations between the different processes on the same content. The most direct method of testing this hypothesis is to submit the intercorrelation matrix of all process scores, where each process-content combination is regarded as a test variable, to circumplex analysis (Guttman, 1954).

The uniform, perfect, additive circumplex is described as follows:

If a uniform circumplex exists, then

$$t_{ji} = \begin{cases} c_{ji} + c_{j+1,i} + \dots + c_{j+m-1,i} & (j \leq n-m+1) \\ c_{1i} + c_{2i} + \dots + c_{j-n+m-1,i} + (c_{ji} + \dots + c_{ni}) & (j > n-m+1) \end{cases}$$

where t_{ji} represents a test (t_j) with elementary additive components (c_i). If we assume that all elementary components to be uncorrelated, i.e.,

$$r_{c_p c_p} = 0 \quad (p \neq q)$$

and consider the special case where all elementary components have equal variance and where $m \geq n/2$ then:

$$r_{jk} = \begin{cases} 1 - \dfrac{k-j}{m} & 0 \leq k-j < n-m \\ 1 - \dfrac{n-k+j}{m} & n-m \leq k-j < n \end{cases}$$

A matrix of intercorrelations which satisfies these restrictions will have two unique characteristics; the column totals will be equal, and each row of the table will have the same entries as the preceding row, but moved one space to the right, the end one moving to the beginning. In the case where the circumplex is not equally-spaced, then the column

totals will not be equal. The matrix can then no longer be considered a circulant. It is possible to define a quasi-circumplex in the same manner as a quasi-simplex is defined.*

Psychological Structure

The third specific problem is to explore the psychological structure of the taxonomy. The reader will recall that three criteria in constructing the taxonomy were educational meaningfulness, logical rigor, and psychological meaningfulness, in that order of priority. It is the latter characteristic upon which this specific aspect of the study focuses.

The major objectives were (a) to determine the kinds of cognitive aptitudes, as defined by tests in the Kit of Reference Tests (French, Ekstrom and Price, 1963), and their regression weights, which are related to scores attained on tests of each level of the taxonomy, (b) to determine if these equations for a given level, but calculated over different

---

* Parenthetically there appears to be an inconsistency in the taxonomy. On the one hand, the behaviors are regarded as being general over content. On the other hand, the Knowledge category is seemingly defined as the content on which the processes are to be exercised. Strictly speaking it would seem that the Knowledge category should be defined simply as the process of recall or remembering. Unfortunately, if one does so, then he encounters difficulties perhaps more serious than those posed by the seeming inconsistency. Specifically, it seems that the measurement of recall cannot be separated from the measurement of content. For example, if one asks, "In what year did Columbus discover America?" and he receives no response, then one cannot with assurance ascribe the failure to either the student's not having stored that bit of information or his inability to recall it from storage if in fact it is stored. However, if the item is changed to "In what year between 1400 and 1500 did Columbus discover America?" and the student responds "1492", then one would probably grant that the relevant content was in storage and the process of recall was adequate. But it should be noted that the addition of "between 1400 and 1500" altered the stimulus value of the item. Thus, a question is raised about what the Knowledge category of the taxonomy should be. Should it be "stored content" as it seems to be at present or should it be "the relative accessibility of stored content"? The latter would seem to be more consistent with the other major categories of the taxonomy. At any rate, it poses an interesting research question. Because of the question raised here, this study was structured so that this point would not be a crucial one; nevertheless, it deserves attention.

contents, are similar or dissimilar, (c) to determine the changes in factors and weights in the equations for each of the several levels of the taxonomy, for different contents separately, and (d) to determine the similarity of equations for the same level of the taxonomy but calculated on different grade-groups of students.

Jointly these studies should reveal the structure of the taxonomy in terms of more elemental cognitive aptitudes, whether process transcends content, whether the structure of the taxonomy as defined by these regression equations is a function of the grade-level of students or is independent of it, and whether there are systematic changes in the regression equations from level to level of the taxonomy.

C. Significance of the Problem

The taxonomy is receiving increasingly wider use in education. The current uses go far beyond merely facilitating communication about educational objectives. These include establishing series of educational objectives which are ordered ostensibly according to increasing complexity, the analyses of educational behaviors into the taxonomic classification, exploratory experimentation in the development of instructional programs intended to foster cognitive development, etc. Although these uses constitute justifiable commonsense employment of the taxonomy, it does seem that their legitimacy rests on the assumption that the taxonomy structure is verifiable. It is the intention of this project to initiate these validational studies.

It should be acknowledged immediately that we have a serious reservation about the propriety of our effort to validate the taxonomy. The developers of the taxonomy state that the phenomena to be classified are "intended student behaviors," i.e., the behaviors which an item is intended to evoke. They state that the phenomena are not the behaviors which students actually execute when responding to a test item. If one holds rigorously to their stipulation, then the validity of the taxonomy rests on the agreement among raters in designating the behaviors which an item is intended to evoke. Therefore, the actual student behaviors are irrelevant in the validation process. We acknowledge the distinction between intended behaviors and actual behaviors which they drew, but have chosen to proceed on the assumption that the usefulness of the taxonomy in educational planning and research depends upon the empirical demonstration that the structure of the taxonomy portrays actual student behaviors. Consequently, the analyses presented herein and the interpretation of the

results of them should be considered as relevant to our re-definition of the phenomena to be classified. Thus, the results do not necessarily bear on the validity of the taxonomy when considered in terms of the announced purpose of the taxonomists.

The validation of the taxonomy would have several distinct theoretical and practical advantages. First, it would enable those who use it in the intended sense to do so with greater confidence. Second, it would provide a sound basis on which to develop sequential educational programs intended for the systematic development of increasingly more complex objectives. Third, if the processes are found to be general over content, then the processes themselves might be considered as desirable educational objectives for no other reason than their pervasive and enduring nature and the temporality of most content. Fourth, it would provide a novel theoretical framework in which to consider certain educational problems; e.g., educational underachievement might become amenable to study as a phenomenon which is explainable on the basis of cognitive disability or under-development of a nonorganic and nonpathological kind, instead of on the prevalent but somewhat secondary ground of invoking social and emotional variables as the primary explanatory concepts. Fifth, if it can be demonstrated that the major levels of the taxonomy are unique arrangements of known aptitudes which pervade all levels or if it can be shown that each succeeding level is explainable by an increasingly larger set of aptitudes, then a new research avenue will be opened which might lead to the deliberate development of broad aptitudes or processes and perhaps the experimental creation, through novel instructional programs, of aptitudes which have not yet been identified.

II. BACKGROUND

This section consists of two major parts. The first deals with related studies which the investigators conducted immediately prior to the study reported herein. The second deals with a survey of the reported research on the taxonomy. These are presented separately because most of the results of earlier studies by the authors do not appear in the literature and because the present study is an outgrowth of them.

A. Earlier Studies by the Writers

In 1962, the writers received a grant-in-aid from the Graduate Council of the Florida State University for "exploratory studies in the validation of measures of cognitive processes." Two kinds of studies were undertaken initially; to determine whether test items could be classified reliably according to the major levels of the taxonomy and to determine the characteristics of test data when partitioned according to levels of the taxonomy. As a consequence of them, a third study, the construction and preliminary try-out of taxonomy-type tests, was undertaken. Each is described below.

The first two studies, inter-judge reliability of item classification and statistical characteristics of taxonomy scores, centered on the Reading Test and the Arithmetic Computation Test of the advanced battery of the Metropolitan Achievement Tests. The Reading Test consists of seven short reading passages each of which is followed by four to nine four-choice items. There are forty-four items in the whole test. The Arithmetic Computation Test consists of forty-five items dealing with subtraction, multiplication, and division of mixed, whole, and decimal numbers; calculation of interest rates, volumes, areas, percentages; calculations based on data presented in graphs, pie charts, etc. Each item is three-choice in the test booklet but the answer sheet provides two additional responses; DK, "don't know," and NG, "not given."

Classification of Items by Judges

Reading test. The forty-five items of the Reading Test were sorted according to the six major levels of the taxonomy by two sets of judges. All judges were familiar

with the taxonomy. Seven of the eight judges had at least three years of graduate instruction in tests and measurements and the other judge was a third-year doctoral student who minored in tests and measurements. The judges were asked to classify each item according to the major levels of the taxonomy. They worked independently and each had available a copy of the Taxonomy and a test booklet which contained the items and reading passages. A set of five of the judges sorted the items initially, then four months later a second set--three of the original judges and three additional ones--classified the same items.

When the modal classification of an item by each set of judges is considered, then only five of the forty-four items received different modal classifications by the two sets of judges; e.g., the modal judgment of one set was Knowledge, and the modal judgment of the other was Comprehension; twenty-four items received the same modal classification by both sets of judges; and five were nearly similar in that one set reached a bimodal classification whereas the other set reached a modal classification which corresponded to one of the levels in the others' bimodal classification. However, it should be noted that three of the judges were common to both sets and four months elapsed between the two sortings.

The first set of judges gave modal classification to items as follows: Knowledge, ten items; and Comprehension, twenty-nine items. They reached the following bimodal classifications: Knowledge-Comprehension, two items; Comprehension-Analysis, two items; and Comprehension-Synthesis, one item. The second set of judges classified modally ten items as Knowledge and twenty-two items as Comprehension, and eleven items were classified bimodally as Knowledge-Comprehension.

The extent to which judges spread their ratings over items is as follows. The first set of judges used only one or two levels of the taxonomy to classify thirty-one items, and only three or four levels to classify the remaining thirteen items; however, only one item was classified in four levels. The second set of judges used one or two categories or levels with thirty-six items and three categories with the remaining eight items.

The three judges who were common to both sets of judges gave exactly the same rating at both times to only eight items; i.e., they agreed completely with each other at one time, then agreed completely the second time, and the ratings at both times were identical.

For eighteen items, each judge gave the same rating at both times; for twenty-one items, only two judges gave the same rating at both times, and on five items just one of them gave the same rating at both times.

The phenomena on which the taxonomy is based is the behavior which an item is intended to elicit. The description of the agreements and disagreements among judges which are presented above suggest that there is substantially less agreement among them than one would expect. To a great extent their disagreements can be rationalized on three bases which were revealed in group discussions with them. First, they did not use the same frame of reference. Second, apparently some of the items can be solved by several processes each of which appears at a different level of the taxonomy. Third, the limited range of cognitive process represented in the Reading Test probably speciously magnified differences among them.

The failure of judges to share a common frame of reference might be illustrated by the following item. "Which of the following words is a synonym for the word 'lazy'"? Some judges might classify the item as Knowledge because they believe that the typical student for whom the test is intended already knows many synonyms for that word and can respond on that basis; therefore the test author intended it to be answered by recall; i.e., Knowledge. Other judges might classify the item as Comprehension on the basis that the test author intended the synonym to be chosen on the basis of contextual clues in the reading passage. Many such differences in the frame of reference were uncovered during discussion with judges. The fundamental point is that the judges differed in their assumption about the prior knowledge which a student brings to the test situation. In turn each assumed that the test author held his frame of reference and as a consequence judges differed with regard to their opinions about the behaviors an item was intended to evoke. Concerted work with the judges did not cause them to narrow their different conceptions of the prior knowledge which students bring to the test situation. Apparently, there is such a diversity in the educational experiences offered to students at the same grade level but who attend different schools that each judge could argue with conviction about the tenability of his frame of reference on the basis of his direct experience from teaching in the public schools.

It also became clear that their ratings differed because they believed some items could be answered correctly on the basis of processes which appear at

different levels in the taxonomy. One kind of item required students to select the most appropriate title from a list of titles for a reading passage. One judge classified the item as Comprehension because he regarded the task as simply translating and reducing the reading passage into a title. Another judge classified the item as Analysis because he regarded the process as a line-by-line analysis of the passage to determine relationships among them that would serve as a basis for choosing a title. Disagreements on some other items could be attributed to one judge regarding the problem solving strategy as the manipulation of content in the passage to reach an answer and then to locate it in the list of alternatives, whereas another judge regarded the strategy as the systematic elimination of distractors from the list of alternatives. Parenthetically, we subsequently presented the same item stem with and without alternatives to students who were asked to verbalize their solutions and did determine that the actual problem solving strategies did differ in many cases.

In summary, judges exhibited less agreement in their ratings than we had expected. Many of these disagreements were due to their different assumptions about the prior experience of students and about the problem solving strategies students would use. The intra-judgment agreement was greater than the inter-judge agreement which simultaneously suggests that the judges were consistent in their different assumptions and that, within a set of pre-conceptions on the part of the judge, the taxonomy provided a rather unambiguous guide for classifying the items.

Arithmetic Computation test. Seven judges classified each of the forty-five items of the Arithmetic Computation test. Their modal ratings were: Knowledge, twenty-five items; Comprehension, four items; and Analysis, one item. Six items received bimodal classifications: Knowledge-Application, four; Comprehension-Application, one; and Application-Synthesis, one. No item was rated identically by all seven judges and only each of six items received identical ratings from six of the seven judges.

Analyses of the ratings revealed considerable intra-judge consistency over similar kinds of items. Group discussion revealed again that differences among them was due mainly to the different assumptions about the relevant experience which students bring to the test situation. Generally, judges who had the most experience

teaching students of the age the test is intended were more likely to rate items at lower levels in the taxonomy than were judges who had less teaching experience. Other differences among judges were noted but these reduce to imputed differences in problem solving strategies which were discussed in the preceding section.

Analysis of Data from Tests

The material presented in this section is based on three hundred students for whom reading, arithmetic, and scholastic ability scores were available. The students were ninth-graders in a single public school system. Three groups of 100 subjects each were selected on the basis of their scores on the Verbal Section of the School and College Ability Test so that the groups were statistically non-overlapping. The high group was selected randomly from students who scored more than one standard deviation above the mean; the average group was selected randomly from all students who scored at or within one score point of the mean of the distribution; and the low group was selected randomly from all students who scored more than one standard deviation below the mean.

Item data. On the Reading Test, eleven items were classified as Knowledge and the remaining items as Comprehension. The average difficulties on the Knowledge and Comprehension items were .75 and .69, respectively. The average difference between the percentage of high and low ability groups passing an item was .27 on the Knowledge items and .42 on the Comprehension items. Item factor analyses based on exact tetrachoric coefficients were run for the three ability groups, separately. Eight to ten factors emerged. At least one factor in each matrix might have been named "vocabulary" however, not all items which, on logical analysis, would be placed in such a factor actually loaded on it. The majority of the other factors seemed to define the several reading passages which constituted the Reading Test. Attempts to sort each of the items into finer subdivisions of each level of the taxonomy failed to create item groupings which corresponded to the factors which were obtained.

On the Arithmetic Computation Test thirty-two items were classified as Knowledge, six items as Comprehension, five items as Application, and two items

as Analysis. The average item difficulties for these were .79, .72, .58 and .42, respectively. The average item differences between percentage of success of the low and high ability groups were: .20, .31, .48, and .42, respectively. No item factor analyses were run on data from the Arithmetic Computation Test.

Generally, the data from both tests lend support to the imputed hierarchical structure of the taxonomy. The item difficulties increased directly with the increase in classification level of items. The data also supported our hypothesis that the spread between the success of the high and low ability groups would increase as the classification levels of items increased. Although the item factor data failed to reveal factors which could be called Knowledge and Comprehension but instead yielded a "vocabulary" factor and some rather poorly defined factors corresponding to the reading passages, we are disinclined to take the factor analyses seriously because of the large number of items on which there were extreme splits on pass and fail.

<u>Part-score data</u>. After the test items in the Reading Test and the Arithmetic Computation Test had been classified, separately, according to levels of the taxonomy, the answer sheets were scored to yield these taxonomy-type processes scores. The intercorrelation matrix of them appears in Table 1 and the factor matrix appears in Table 2.

The rotated factor matrix reveals that Reading Knowledge and Reading Comprehension define one factor which might be called verbal ability. Arithmetic Knowledge and Arithmetic Application apparently define another. Theoretically, this grouping would not be expected; however, it might partially be explained by conflicting frames of references of the item classifiers. Since the Arithmetic Test did not include a reading passage which explicated the principles which were to be applied during the test, then it follows that students had to apply principles they already knew. Consequently, judges disagreed on whether such items should be regarded as Knowledge or Application. The factor on which Arithmetic Analysis is highly loaded should, in our opinion, be ignored because of the few number of items on which the Analysis score was based.

The inference we drew from these studies was that if one wishes to investigate the validity of the taxonomy then he could not do so with data obtained from typical

TABLE 1

Part-Score Correlation Matrix

Tests	1	2	3	4	5	6
1. Reading Knowledge		.70	.53	.49	.45	.33
2. Reading Comprehension			.54	.60	.56	.41
3. Arithmetic Knowledge				.61	.61	.45
4. Arithmetic Comprehension					.56	.46
5. Arithmetic Application						.56
6. Arithmetic Analysis						

TABLE 2

Oblique Rotation of Factor Analysis
of Part Scores*

Tests	I	II	III	IV
1. Reading Knowledge	.90	.21	.08	.10
2. Reading Comprehension	.79	.18	.19	.31
3. Arithmetic Knowledge	.30	.80	.12	.32
4. Arithmetic Comprehension	.30	.28	.19	.85
5. Arithmetic Application	.27	.70	.47	.12
6. Arithmetic Analysis	.14	.20	.92	.14

*Unity was placed in cells of principal diagonal

standardized achievement tests. These tests did not give equal representation to all levels of the taxonomy. Furthermore one could not assume with confidence that students brought the same content knowledge to the test situation. In short, we concluded that it would be necessary to construct tests according to the Taxonomy in order to obtain the data we needed. The next section briefly describes our initial efforts in this activity.

Construction of Taxonomy-Type Tests

Several specifications were set for the taxonomy-type tests which were to be constructed: (a) each should be suitable for use with secondary school students; (b) each test should contain a reading passage which supplies the content on which the test items are based; (c) items should be distributed approximately equally over the six major levels of the taxonomy; (d) all items should be the multiple-choice type; and (e) test administration time for each should not exceed forty minutes including all pretest and posttest directions.

Three reading passages were chosen on the bases of their probable interest value, appropriate reading level, and probable unfamiliarity to students. One dealt with the periodic table of elements; another with relationships among surface area, cross-section, and weight; and the third with the way of life in time past in Nayon, Equador.

The project staff prepared test items according to taxonomy levels for each reading passage. Each item was criticized and discussed in staff meetings. Primary attention was devoted to classifying appropriately each item according to the taxonomy and editing it to make it technically good.

A test booklet was then prepared consisting of directions, reading passage, and the test items in random order. Three such preliminary versions were made and administered to eighty to 140 students.

On the basis of observations of students during the preliminary testing and analyses of item data, certain revisions of items, preparation of new items, rearrangements of the items in the booklet, and test administration procedures were executed. These are of importance and interest but the discussion of them is appropriately withheld until the next major section of the report which

deals with the construction of tests for the federally supported project.

Considerable difficulty was encountered when revising the Nayon test. The content of the reading passage was not sufficiently dense to enable writing a large number of items about it, nor sufficiently complex to enable preparing items for the upper levels of the taxonomy. Nevertheless, it was revised and re-administered but student reactions to it during its administration and the results of the preliminary analyses of data from it which suggested low reliability, caused us to discard it entirely. Consequently, it is not mentioned hereafter. The following account is essentially the content of an article which was prepared by the writers (Stoker and Kropp, 1954).

The tests. The revised form of the test based on the periodic table of elements, hereafter referred to as AS (atomic structure) consisted of thirty-six multiple-choice items, six for each level of the taxonomy. The revised form of the test on relationships among surface area, cross-section, and weight, hereafter referred to as RS (right size), consisted of thirty-six multiple-choice items and three free response items; which were nearly equally divided among the six levels of the taxonomy.

Rater agreement on items. Five judges independently classified the items of the AS test according to the major levels of the Taxonomy, and four different judges similarly classified the items of the RS test. All judges were doctoral students in educational measurement and were familiar with the Taxonomy.

On the AS test, eleven of the thirty-six items were unanimously classified congruent with the categories whose processes the items were intended to evoke. On nine other items, only one judge deviated from the other four. (Six of the nine disagreements were due to the same judge.) On all but two of the remaining sixteen items, three of the five judges rated each item in the category whose process it was intended to evoke.

The four judges who classified the RS items agreed as follows: for eleven items their agreement was unanimous and perfectly related to the categories the items were intended to evoke; for sixteen other items,

three of the four judges agreed with each other and the intended category; and for the remaining items two of the four judges were in agreement.

Empirical support of taxonomy. Since the taxonomy was designed to be hierarchical and cumulative, the data should exhibit these characteristics: (a) the mean scores on the levels should decrease as the complexity of the levels increases; (b) the intercorrelation matrix of level scores should reveal simplical structure as described in the Guttman Radex Technique; and (c) factor analyses should reveal a structure consisting of a unique factor loading only on the Evaluation subtests, a general factor, and four group factors which might be named Comprehension, Application, Analysis, and Synthesis.

To gather data to test these hypotheses, the two experimental tests were administered to the student bodies of two four-year high schools. Six hundred students were tested from one school and four hundred from the other.

Hypothesized inverse relationship of level and mean score. To examine the hypothesis that mean difficulty increases as complexity of level increases, the ratios mean score to total items was computed for each grade group on both the AS and the RS tests. The ratios appear in Table 3.

The ratios generally support the existence of a hierarchical structure. Mean performance decreases as process complexity increases, thus indicating support of the taxonomy and some construct validity of the experimental tests. The exception to these general statements is that the data for the AS test suggest perhaps that the Evaluation level is misplaced in the taxonomy, or the items for it were poorly constructed.

As a further investigation of this hypothesis a random sample of five hundred subjects was sorted into six groups on the basis of total score on the combined Knowledge and Comprehension subtests of AS. Then for each group, mean scores were computed on the four remaining subtests. The data appear in Table 4.

The tabular data reveal (a) that as combined score on Knowledge and Comprehension increases, then so does performance on the remaining subtests, and (b) that for any given level of combined score, the scores on the remaining subtests decrease as the complexity increases. However, this latter statement does not hold for the Evaluation subtest.

TABLE 3

Mean Performance on Experimental Tests by Process and by Grade*

	Grades--School A				Grades--School B			
	9	10	11	12	9	10	11	12
TEST OF ATOMIC STRUCTURE								
Knowledge	.38	.43	.39	.42	.35	.38	.50	.58
Comprehension	.36	.37	.33	.37	.32	.30	.38	.48
Application	.38	.40	.41	.43	.35	.34	.46	.44
Analysis	.34	.38	.33	.40	.34	.33	.42	.49
Synthesis	.28	.28	.26	.27	.27	.30	.30	.38
Evaluation	.38	.45	.40	.47	.37	.41	.45	.51
TEST ON RIGHT SIZE								
Knowledge	.53	.60	.67	.68	.53	.55	.68	.70
Comprehension	.52	.58	.60	.57	.48	.58	.61	.68
Application	.24	.28	.31	.32	.25	.30	.36	.32
Analysis	.30	.33	.34	.38	.31	.34	.38	.38
Synthesis	.17	.23	.22	.24	.15	.19	.24	.28
Evaluation	.19	.21	.25	.22	.19	.22	.24	.24

*Table values are equal to the mean score for the test divided by the number of items in the test.

TABLE 4

Mean Scores for Higher Taxonomy Levels When Subjects are Grouped on Combined Knowledge and Comprehension Scores (Form AS) (N=500)

Combined Score (Know. & Comp.)	Application	Analysis	Synthesis	Evaluation	N
11-12	4.20	4.00	2.00	3.40	5
9-10	3.22	3.39	1.89	3.61	18
7-8	3.35	2.89	1.94	3.13	62
5-6	2.55	2.07	1.65	2.68	147
3-4	2.22	1.78	1.51	2.28	169
0-2	1.47	1.56	1.58	1.88	99

Presence of simplical structure. According to Guttman's simplex theory, if a group of tests differ only in complexity, then this fact will be observable in a matrix of intercorrelations among the tests. The hypothesis of hierarchical order for the taxonomy levels, then, lends itself to examination by simplex analysis. The theoretical rationale for the simplex was presented in Section II in the context of hypotheses to be tested.

Twenty correlation matrices, arising from the administration of the two experimental forms in both schools, were subjected to simplical analysis. None of the matrices satisfied the requirements of a perfect simplex. However, in approximately one-half of them, the Knowledge, Comprehension, Application and Analysis subtests were ordered correctly but the Evaluation and/or Synthesis subtests were repeatedly out of order.

Factor analyses. Very little evidence of the expected factor structure was found. For several matrices, the factors were identified as content factors; i.e., AS and RS. When factor matrices for different grades were examined two trends were noted: (a) Knowledge and Comprehension on the AS test persistently loaded together

over grade levels, and (b) a definite increase in the size of factor loadings was observed as grade level increased from 9 to 12.

Summary. Generally the study might be summarized as follows. Interjudge agreement was found with respect to cognitive process being sought in items. The data give general support to the hierarchical structure of the Taxonomy. The data suggest a possible misplacement of Evaluation in the structure. Factor analyses failed to support the hypothesized structure of the Taxonomy.

Critique

The results of these early studies of the classification of items and the structure of the taxonomy weakly supported the validity of the taxonomy. As these studies progressed it became clear that those dealing with the agreement among raters' classifications of items were interesting but not central; however, they did document the clarity with which the processes are described in the Taxonomy. The analyses of data from the commercially available tests yielded results which were somewhat helpful in illumining the taxonomy but their overall value was slight because the vast majority of items in the tests were at the lowest level and none of the upper level processes were assessed at all. At this point, it became evident that the most relevant data would be generated by specially prepared tests which would yield appraisals of student behavior at each of the six major levels of the taxonomy. To construct such tests was regarded as a task which would prove to be more time consuming than complex. Eight months of half-time work on the part of six to ten people in writing approximately 500 items and testing about 1,900 students, coupled with about 30 hours of computer analyses of the resulting data produced the unimpressive results which were presented above. The task had been misjudged; it was time consuming and complex.

The staff had specific misgivings about the tests and the data which came from them. The test content was entirely chemistry and physics; every attempt to build a test on social science content was abortive. There was a pronounced lack of confidence in all items which were prepared for Synthesis and Evaluation and there was a lesser, but significant feeling, that items for those levels could not be

written in the multiple-choice format. In fact, we departed from our specification and included free response items in the RS test. The subtest reliabilities were quite low probably due to the shortness of the subtests. It seemed that no more technically good items could be wrung from the reading passages. Furthermore, adding items would cause the overall testing time to become longer than the length of a public school classroom period. Traditional item analyses techniques seemed to be completely inappropriate for evaluating taxonomy-type test items. In short, there were serious reservations about the quality of what had been achieved. Nevertheless, the results gave slight support and this caused the staff to prepare a proposal to be submitted to the U. S. Office of Education which would provide for a more extended and comprehensive study of the taxonomy.

B. Studies Reported in the Literature

Very few studies that deal directly with the Taxonomy appear in the literature. Those which could be located and which are directly relevant to the study are summarized here in three categories; classification of items by judges, statistical analyses of taxonomy-type data, and test banks.

Classification of Items by Judges

McGuire. A series of studies were undertaken at the University of Illinois College of Medicine in which the process approach to constructing and analyzing medical examinations were used. (McGuire, 1963). One of the early studies was to classify the items in the 1961 National Board Examination which was issued by the National Board of Medical Examiners. Four panels of three members each were formed to rate the items. Each panel represented a medical speciality and the three panel members were expert in that speciality. Each item was rated with reference to an eight-level adaptation of the Taxonomy. The levels were: recall, recognition of meaning, selection of relevant generalizations, simple interpretation of data, application of principles to situations of a familiar type, application of principles to situations of an unfamiliar type, evaluation of a total situation, and synthesis. Each panel rated only items in its speciality. Of the 683 items, each of 61 percent of them received identical ratings by the three judges from the relevant panel. On each of 13 percent of the items, two of the

judges rated the item similarly but the third rated it either one level above or below their rating. Only on each of 7 percent of the items did all three judges disagree. More than half of these 7 percent of items were at the lowest three process levels of the taxonomy.

Stanley and Bolton. Stanley and Bolton (1957) reported several related studies of the extent of agreement of item classifications which were made by their graduate students. They administered Terman's Concept Mastery Test, Form T, to 46 students in a beginning graduate course "Principles of Appraisal and Evaluation in Education." The eight highest scoring students (101-135) were asked to classify independently the 227 items in Gerberich's Specimen Objective Test Items, A Guide to Achievement Test Construction according to the six major levels and the twenty-three sublevels of the Taxonomy. Prior to this classification exercise, all class members had studied the Taxonomy for four weeks. The percentage of 8, 7, 6, 5, 4, 3, and 2 agreements on major level-and-sublevel were 5, 13, 17, 17, 21, 19, and 8, respectively. One item was assigned to a different sublevel by all judges. The agreement among the eight judges was quite high when only the six major levels were taken into account. The raters classified the items as follows: Knowledge, 51 percent; Comprehension, 19 percent; Application, 8 percent; Analysis, 5 percent; Synthesis, .6 percent; and Evaluation, 4 percent. The raters judged that 12.5 percent of the items did not fall in the cognitive domain.

Thirty-one items for which there was high agreement were cast into a test and thirty-six of the remaining class members were asked to rate each item according to major level and sublevel. Responses were scored as follows: two points for a major and sublevel agreement; one point for a major level agreement; and no points for an incorrect major level. The resulting scores were correlated with the Concept Mastery Test scores. These coefficients were obtained: CMT Same-Opposite, .49; CMT Analogies, .43; and CMT Total, .52. (The CMT subtests correlated at .72).

In a further study, the eight original raters and the three students having the next highest CMT scores (91-98) were asked to classify specimen items from the Graduate Record Examinations. Forty-four items were used in the study and these were taken from two prospectuses of the GRE which were distributed

about 1954 by the Educational Testing Service. The major levels assigned to the achievement items were: Knowledge, 49 percent; Comprehension, 21 percent; Application, 4 percent; Analysis, 17 percent; Synthesis, none; and Evaluation, 9 percent.

Summary. Both the McGuire and the Stanley and Bolton studies yielded results quite similar to those of our studies. Trained raters can classify test items into taxonomy categories with high agreement when the basis of the classification is the behavior which the item is intended to evoke. However, neither of the studies nor ours deals with the relationship between the process the author intended to evoke with the item and the process the item actually evokes from the student.

Analyses of Taxonomy-type Data

Anderson. Anderson (1964) constructed a taxonomy-type test and compared scores attained on the first four levels by students in CHEM study and a traditional chemistry course. The taxonomy-type test consisted of six Knowledge items, nine Comprehension items, fifteen Application items and fifteen Analysis items. The test was administered as a pretest in October and as a posttest in the following April. The scores for all levels for both groups were significantly higher in April. No significant differences were found at any level between the two groups on the posttest. When subjects were grouped according to scores on the School and College Ability Test, the only significant difference was found with the Analysis items for low ability students. Factor analyses of the pretest and posttest scores revealed two major factors: Knowledge and Comprehension items defined one factor and Application and Analysis items defined the other.

Dunn and Turner. Dunn and Turner (1959) studied physics achievement in the Australian secondary school system using a modified version of the Taxonomy as a model for test construction. Three measures were obtained from a total of 1,526 students who were enrolled in physics. These measures were a prematriculation examination, a matriculation examination, and teachers' estimates of the student's ability in physics. The prematriculation examination was constructed according to a three-level modification of the Taxonomy. These levels were knowledge, application, and understanding. The knowledge

and application levels approximately correspond to those same named levels of the Taxonomy. The subtest of understanding encompassed the other four levels of the Taxonomy. The split-half reliabilities of the prematriculation examination were .837 knowledge (51 items), .828 application (24 items), .888 understanging (55 items), and .933 total test (130 items).

The bivariate distribution of the level scores were triangular indicating that high scores on lower level measures could be associated with either high or low scores on higher level sub-scores. However, a low score on the lower level would not be associated with a high score on higher level subtest. Correlation coefficients of .67 between knowledge and application, .66 between knowledge and understanding, and .71 between application and understanding were found for the three subtests.

Correlation coefficients between the three major variables were estimated from the distribution of the rank order correlations found within individual classes in the sample. The average rank order correlations between each pair of the three variables were .68 between prematriculation and matriculation, .68 between prematriculation and teachers' ranking, and .70 between matriculation and teachers' ranking.

McFall and Chambers. McFall and Chambers (no date) constructed an achievement test, composed of two subtests, on the relationship between plants and animals. One subtest contained recall (Knowledge) items and the other contained problem solving items (higher level cognitive processes). The test was administered to 457 students in seventeen general science classes from grades seven to eleven. The test scores were correlated with first semester science grades, total scores on the Stanford Achievement Test, and scores on the Science section of the Stanford Achievement Test. The investigators hypothesized that the recall items would correlate more highly with achievement measures than would the problem solving items. They found higher relationships between the recall subtest and criterion measures (Stanford Achievement, .44; Stanford Science, .45; and science grade, .41) than between the problem solving subtest and these criteria--.34, .39, and .19, respectively. The differences between each of the three pairs of correlations, e.g., recall and science grade, and problem solving and science grade, was statistically significant.

McGuire. The eight-levels taxonomy which was constructed by the College of Medicine of the University of Illinois was subsequently reduced to six major levels and was named the Taxonomy of Intellectual Processes (McGuire, 1963). The six levels are described as follows: Level 1, recall and recognition of meaning; Level 2, generalization; Level 3, problem-solving of a familiar type which includes (3.1) simple interpretation of data and (3.2) application of a standard combination of principles; Level 4, problem-solving of an unfamiliar type which includes (4.1) analyses of data and (4.2) application of a unique combination of principles; Level 5 evaluation; and Level 6 synthesis.

This taxonomy was used as a framework within which to construct comprehensive examinations. Analyses of the results of those examinations supported the imputed hierarchical and cumulative characteristic of that taxonomy. The hierarchical nature is supported by the intercorrelations between scores from the six levels. There is a tendency for the adjacent levels to be more highly correlated than widely separated levels. The correlation and factor matrices appear in Tables 5 and 6.

Smith. Smith (1965) investigated the scalability of the Knowledge and Comprehension levels of the Taxonomy. A fifty-five item, four-choice multiple-choice test was constructed on five educational psychology concepts. One item was prepared for each concept for each of eight Knowledge sublevels and the three Comprehension sublevels. The test was administered to 341 educational psychology students who had been introduced to the content through lecture of textbook. Because it is claimed that the Taxonomy is arranged from simple to complex and from concrete to abstract, item difficulties should increase as Taxonomy level increases. Smith also claimed that this relationship should hold for the sublevels of any major Taxonomy level. He found a rank order correlation of .76 between the actual ranking of sublevel difficulties and the theoretical rankings. It indicates a general increase in item difficulty as level increases within Knowledge and Comprehension. The intercorrelation matrix of Knowledge and Comprehension sublevel scores failed to form a simplex. Thus the hypothesis of a hierarchical structure of the sublevels of the first two levels was not supported. The inter-item correlations in the three sublevels of Comprehension across the five educational psychology concepts utilized were positive, but no larger than might be expected by chance. Thus, it would appear

TABLE 5

Correlation Matrix

Taxonomy Level	1.0	2.0	3.1	3.2	4.1	4.2	5.0	6.0
Level 1.0: Recall	1.000							
Level 2.0: Generalization	.393	1.000						
Level 3.1: Interp. of Data	.406	[illegible]	1.000					
Level 3.2: App. of Fam. Prin.	[illegible]	[illegible]	.442	1.000				
Level 4.1: Analysis	.280	.415	.462	.420	1.000			
Level 4.2: App. of Unfam. Princ.	.185	.461	.357	.313	.401	1.000		
Level 5.0: Evaluation	.242	.313	.305	.367	.393	.323	1.000	
Level 6.0: Synthesis	.177	.277	.272	.137	.335	.242	.173	1.000

TABLE 6

Rotated Factor Matrix

Taxonomy Level	Factor Number I	II	III	IV	V
Level 1.0: Recall	.922	.071	.155	.057	.139
Level 2.0: Generalization	.437	.142	.004	.642	.327
Level 3.1: Interp. of Data	.425	.197	-.056	.354	.591
Level 3.2: App. of Fam. Prin.	.133	-.077	.199	.105	.856
Level 4.1: Analysis	.036	.406	.277	.246	.614
Level 4.2: App. of Unfam. Princ.	-.022	.102	.219	.891	.137
Level 5.0: Evaluation	.142	.075	.916	.171	.214
Level 6.0: Synthesis	.101	.950	.053	.117	.057

"Level 1.0: Recall has a higher loading on Factor 1, and only on Factor 1; in addition this same factor makes a substantial contribution to scores at the neighboring levels 2 (Generalization) and 3.1 (Interpretation of Data), and makes virtually no contribution to scores at Level 3.2 (Application of Familiar Principles) and above. Other groups of variables which have significant loadings on a common factor are:

Variables	Loadings	Factor
Level 6 and Level 4.1	.95 and .40	on Factor 2
Level 4.2 and Level 2.0	.89 and .64	on Factor 4
Level 3.1, 3.2 and 4.1	.59, .85, .61	on Factor 5

Level 5: Evaluation stands alone as the only variable with a high loading on Factor 3 and it loads significantly only on that factor. These groupings are compatible with the logical structure of the Taxonomy (of Intellectual Processes)."

that the processes measured by the items of sublevels are not general over content.

Thomas. Thomas (1965) studied the three lower levels of the Taxonomy. Her major purposes were to test the cumulative hypothesis of the Taxonomy; to determine the relationship between taxonomy-type test performance and intelligence, and reasoning ability; and to determine if students whose scores were not consistent with the cumulative hypothesis differ in intelligence and reasoning ability from students whose scores are consistent. Data concerning achievement on a taxonomy-type test were gathered utilizing specially constructed tests. Intelligence scores were obtained from the California Test of Mental Maturity and reasoning ability was measured by the Cornell Conditional-Reasoning and Class-Reasoning Tests. The subjects consisted of four sections of seventh grade science students in an urban junior high school. The taxonomy-type tests were developed with one item at each of the three levels for each of eleven generalizations about heat and light. To assure control of content, the experimenter taught each of the sections a one week unit on heat and a one week unit on light. The eleven generalizations were presented and the class discussion was controlled so that the novelty of the higher level items would be preserved. The last day of each week was devoted to testing for the generalizations presented and for gathering the reasoning ability measures.

Guttman's simplex analysis was used to test the cumulativeness of the levels. The correlation matrices of the three levels showed the general pattern of a simplex. The correlations between level scores and CTMM are slightly higher for Knowledge (.41) than for Comprehension (.33) and Application (.32). The opposite was found between level scores and reasoning test scores with the correlations being Knowledge. 46, Comprehension .55, and Application .62. No significant differences were found between the mean CTMM scores for subjects whose level scores were consistent with the cumulative hypothesis and those whose level scores were inconsistent.

Summary. The studies reported in this sub-section give general, but not unanimous, support to the cumulatively hierarchical structure of the Taxonomy. The most ambitious of these studies was conducted by McGuire and it focused on an abridgement of the Taxonomy. The most apparent difference between her taxonomy and the parent

one is the reversal of the Synthesis and Evaluation levels. Consequently, the empirical support she found for her taxonomy partially discredits the parent taxonomy. The Smith study is the only one which dealt directly with sublevels of the Taxonomy. His results do not support the hierarchical structure of sublevels but the results of additional and more complex studies are needed before a definitive judgment can be made. The Thomas study suggests that correlations between taxonomy-level scores and group intelligence scores decrease as the level increases and that correlations between taxonomy-level scores and reasoning ability scores increase as level increases. Similar relationships were found by the writers in their preliminary studies.

Test Bank

The San Diego County Secondary Curriculum Council (Kellogg, 1964) developed an American History Test Bank as a part of an in-service training program. The test bank consists of items which were developed by secondary school teacher of American History and includes items from eight areas of American History. The items were constructed according to the Taxonomy. Inspection of the items in the Test Bank reveals that the higher the Taxonomy level the higher the proportion of free response items. The Synthesis and Evaluation levels are exclusively free response. The Test Bank provides a pool of items from which classroom examinations could be constructed.

This report about the data bank and the Stanley and Bolton study which was presented earlier suggest that the vast majority of test items in the cognitive domain are intended, wittingly or unwittingly, to assess the processes appearing at the lowest levels of the Taxonomy. The writers' analysis of items from standardized achievement tests corroborates this observation.

******

NOTE

Subsequent to preparing this section of the report, Dr. John M. Gordon, University of Hawaii, and Dr. Richard C. Cox, University of Pittsburgh, jointly distributed an annotated bibliography of all published research on the Taxonomy and an annotated list of all Taxonomy research projects known to be currently underway. Copies might be obtained from Dr. Cox.

PRECEDING PAGE BLANK-NOT FILMED

III. TEST DEVELOPMENT

A. Rationale

It was apparent that to determine the validity of the taxonomy, the generality of processes over content and the psychological structure of the processes, it would be necessary to construct tests specifically for these purposes. Analyses of currently available achievement tests revealed that not all of the taxonomy levels were represented and the majority of items appeared to be at the lower three levels.

It was also clear that the appraisal of the processes could be done only if the students had substantial and relatively equal mastery of the content on which the items were based. If they possessed this mastery of content, then a greater proportion of variability of scores could be ascribed to their differential command of the processes. Thus a question arose as to the manner in which equal content mastery might be achieved. Three possibilities were considered. First, to select a content which obviously had been already overlearned by the students. Second, to instruct subjects in a content until they reached a predetermined level of mastery prior to administering taxonomy-type tests. Third, to select an unfamiliar content which appeared to be within their power of mastery, to expose them to it at the beginning of the test, and to permit them to refer to it during the entire testing session.

The selection of a response measure also posed a problem. The alternatives were to have students verbalize their solutions to the items, to record them, then to analyze them for the presence or absence of the process in question; or to collect their forced-choice or free-responses to items, to score them right or wrong on the basis of a logical key, and then to infer that the desired process was used if the student responded correctly, but not otherwise.

Because one purpose of the study was to determine the generality of processes, it would be necessary to select several contents on which to build the tests.

The decisions which were made about these and other problems are presented subsequently.

B. Test Specifications

The specifications on which the preliminary taxonomy-type tests were constructed were presented in the preceding

sections as; the tests should be suitable for use with secondary school students, should contain a reading passage which supplies the content on which the test items are based, items should be distributed approximately equally over the six major levels, items should be multiple-choice, and test administration time for each test should not exceed forty minutes including all pretest and posttest directions.

The preliminary forms of the tests were constructed roughly to those specifications. Observations of students during the time they were taking these tests, comments which they made about them, and results of analyses of data from the tests caused some of these specifications to be modified and others to be introduced.

As part of the Office of Education supported study, additional tests were constructed on the basis of these modified and expanded sets of modifications. Pilot try-outs of them resulted, again, in showing the need to alter the specifications. Rather than to present each successive set of them, the final set is presented here and a discussion accompanies each about the need for the specification and its evolution, if appropriate.

Intended Audience

The decision that the test should be designed for use with secondary school students, grades nine through twelve inclusive, was made at the outset and held throughout. Data from preliminary and final forms of the test reveal that ninth-grade students find the tests to be quite difficult. Preparing a set of items which would assure some success for them and simultaneously provide ceiling for the twelfth-grade students proved to be a difficult task.

Content

To appraise processes it would be necessary to base items on a content on which students had relatively the same mastery so that variability in scores would reflect difference in level of mastery of the processes. Of the three methods which were considered for doing this--select a previously overlearned content, teach the content to a given level of mastery prior to administering the test, and to provide content with the test--we chose to provide content with the test. We could not identify a content which had been overlearned by all prospective testees which also would be sufficiently complex to enable preparing items for the upper levels of the taxonomy.

We discarded the possibility of training students to a given level of mastery in a content prior to the test primarily because it would have been too time consuming. Data were needed from several thousand students so the estimated time that would be needed to train them in groups of thirty made this possibility completely untenable. However, we continue to believe that doing so will yield the most valid data if costs in time and money are unimportant.

The content would have to be relatively short, within the reading ability of students, within the mastery of students, of sufficient internal complexity to enable writing items at all levels, and equally unfamiliar to all students at the outset. Particular problems which were encountered in selecting and abridging such passages are described in later sections.

The relevant content was presented as a reading passage at the front of each test booklet and students were directed to read it before attempting to answer the items and they were free to refer to it anytime during the test. Their rather disappointing test performance caused us to abandon the random order in which items were presented. All Knowledge items were presented initially and they were ordered according to the appearance of answers in the reading passage with the consequence that responding to the intial set of items caused the student to read through the initial passage again.

Distribution of Items over Levels

The specification that the number of items should be constant over six levels was met in many of the preliminary forms but was ultimately abridged. Each of the final forms contains twenty items for each of Knowledge, Comprehension, Application and Analysis, five items for Synthesis, and ten items for Evaluation. The scoring scheme provides the same maximum score for all levels. The departure from having equal numbers of items at all levels was prompted by our becoming convinced that Synthesis and Evaluation items could not be written in the multiple-choice format, which was also a specification. Quite abruptly we discarded that format for those levels and substituted free-response items.

Item Format

As stated above, this specification was altered. The first four levels of the taxonomy are assessed by multiple-choice items, and the upper two levels by free-response items.

Administration Time

The initial specification was that an entire test, including all preliminary and final directions, should be administered within a class period. This requirement was held through many of the preliminary versions but was finally abandoned as being unrealistic. Working within the confines of a single class period and the restriction that equal numbers of items appear at each level, resulted in producing a test which contained only six items per level. The brevity of such subtests resulted in unacceptably low reliabilities. It should be noted in this context that as the complexity of an item increases, the solution time does too. The amount of time which students spent on Analysis, Synthesis, and Evaluation items was substantially greater than our expectation. Consequently, new items were added and time limits were lengthened. The final forms are designed to be administered in two sittings of fifty to sixty minutes each. The first sitting includes the items for the first four levels and the second sitting consists of the Synthesis and Evaluation items. The parts must be administered in that order.

Test Forms

Our initial specification was to build several forms each based on a different content. This specification was held throughout but more in spirit than in fact. The initial decision was to build two tests, one based on science content and the other on social science content, and an equivalent form for each. Quite quickly it became clear that equivalent forms could not be developed so four forms were constructed, two on science content and two on social science content. The extent to which each is equivalent to the other three is essentially a restatement of the question of whether the processes transcend content and it is dealt with in the section on results.

C. Selection of Passages

The content for the tests was to be supplied in a reading passage which would precede the items in the test booklet. The following requirements were established for the passage; 600-900 words in length, probable high interest value to secondary school students, social science and science content, within the reading ability of secondary school students--actually the difficulty level was set to be about the ninth-grade level--and the content for the passages should be selected from published materials rather than to be written by the staff.

Generally the following kinds of publications were searched for passages; professional journals, publications which carried popularizations of science and social science articles, and books which consisted of extended accounts of science or social science topics which were presented in layman's language.

Articles were gathered which dealt with the following topics; science of soap bubbles; evolution of a river; evolution of a star; Islam religion; interrelationships of surface area, cross-section, and weight; life in a primitive village; stages of economic growth; atomic structure, the effect of an earthquake on a country; the nature of glaciers; and relationships between architecture and climate.

The reader should note that the selection of these articles and the eventual discarding of some of them were not executed in the orderly fashion implied by this account. First, the project staff divided into teams to select articles, to present them to the entire staff for their approval, and then to prepare items based on the passage. Second, the activities among subgroups were carried out in the same sequence but the stage of development was not coordinated over the groups. Furthermore, discarding tentatively selected passages occurred at nearly every stage of the process ranging from the time the passage was presented to the group for the first time to a point after which the resulting test had been administered in two preliminary versions to at least five hundred or more students. The balance of this topic is devoted to difficulties which the staff encountered with the passages, student reactions to the passages, abridgements of the passages, and particular reasons for discarding some of them.

Some articles had to be discarded because they simply did not contain enough complex materials or enough principles to permit writing a large number of items about them. If articles did not contain stated principles, scientific laws, or generalizations then the articles were eventually discarded because it became extremely difficult to write anything but Knowledge and Comprehension items about them. This realization came quite late during the initial period of test construction. Sometimes hundreds of hours were invested in preparing test items before it became clear that the intellectual ceiling of the passages was too low.

One passage was carried through two preliminary test administrations before it was abandoned. Although the passage seemed to fit exactly our needs, we later learned from student comments and item analyses of their responses that the passage

was probably too complex for them to grasp in the short time that was allowed at the beginning of the tests. Between administering the first and second versions of it, the passage was made more redundant but the resulting passage was too long to be suitable.

Another passage was carried through two field try-outs but was abandoned because it apparently bored students to the point where they did not want to work with it. This reaction was noted more frequently with social science materials than science materials.

Still another passage had to be abandoned because, at the time the preliminary version was nearing completion, new instructional materials were introduced in the schools which virtually overlapped the test content. Although we believed the test would serve our purpose, we discarded it because the content would soon be relatively familiar to all secondary school students and as a consequence the test would be virtually useless to anyone who wished to use it in the future.

The four passages which were finally adopted dealt with the Lisbon earthquake, evolution of glaciers, a theoretical account of stages of economic development, and atomic structure. Only one of these, atomic structure, was a passage which the staff had worked with from the time the original study began.

The four passages were used in their original forms at the outset but experience with them through item writing and analysis of student responses showed that each had to be altered in some manner. It was necessary to introduce additional content into each to enable writing Analysis, Synthesis, and Evaluation items which were acceptable to us. The passage about the Lisbon earthquake had to be shored up by weaving in various philosophical theories about why earthquakes occurred. The passages on glaciers was supplemented with several tables and illustrations which we found to be a most useful technique for introducing a considerable amount of content without taking an inordinate amount of reading time from students. The passage on atomic structure was expanded to include a Twentieth Century counterpart of Mendeleef's original table.

The four test booklets which contain these reading passages appear in Appendices A through D, inclusive.

D. Preparation and Classification of Items

The project staff divided into teams each of which was responsible for preparing and classifying items for a

reading passage. Each team consisted of a content specialist who had had experience teaching public school students in the general content area, a measurement specialist, and usually a person who combined both the experience and training.

The initial requirement laid on these teams was to prepare a certain number of each items at each level of taxonomy. The teams were most likely to prepare items at the lowest levels first, then to move into the higher levels only to find that the reading passage was too lean for them to do so. The requirement was then changed to preparing items on a round-robin basis; i.e., to write an item for each level and then to repeat the cycle. This procedure proved to be much more economical because deficient passages were identified much quicker.

These writing experiences led to the observation that the analysis of passages for adequacy and the facilitation of item writing would be considerably enhanced if at the outset the reading passage was analyzed for its internal theoretical and logical structure. Such an outline of the structure of the passage was useful in inferring whether upper-level items could be written and subsequently, it proved to be valuable in determining how to augment the passages with additional content. With the outline, it was relatively easy to determine the characteristics which supplementary material should have in order to increase the complexity of the passage, broaden the knowledge basis of it, etc. Augmenting passages without regard to having a clear notion of the internal structure often resulted in merely adding a layer of content which extended the breadth of the passage but rarely extended its depth.

Each writing team presented on a regular basis its items to the entire group for review and criticism. Prior to the group meetings all members received a copy of the reading passage and a list of items. Each item carried a code containing the classification of the item according to the taxonomy and the name of the item writer. Each item was considered individually at the group session. The general sequence of consideration was as follows.

First, the participants, except for those who prepared the item, attempted to establish the level of the taxonomy which the item was intended to represent. No attention was given at this time to classifying an item more precisely than by major levels. However, item writers deliberately attempted to sample each of the subdivisions within each level.

Second, when a difference between the writer's intended classification and the group's classification occurred, it was

reconciled whenever possible. When an agreement could not be reached the item was discarded. Principal sources of difficulty were Synthesis and Evaluation items and any other items for which responses might have been learned by students prior to their exposure to the test. Differences among the project staff about Evaluation and Synthesis items were difficult to reconcile during the time when all items were written in the multiple-choice format. These disagreements almost entirely disappeared when the free-response format was adopted. When disagreements arose which were basically differences in attributions about the experiential background which a student would bring to the test, the disagreement was hard to resolve except arbitrarily. Prolonged and fruitless debate caused us to adopt a resolving principle; the group would bow to the judgment of the staff member who had previously taught in that content area at the secondary school level.

Third, after a classification of the item had been agreed on, the group criticized the item. The stem and alternatives were considered in turn. If only editorial changes were made, then the item was accepted for use. If changes in the substance of the stem or alternatives were made, then the group reconsidered the classification of the item. Oftentimes seemingly slight changes in the item caused a change in its level. This general working scheme was followed weekly in group meetings until all items for a trial form of a test had been accepted.

As stated earlier all participants had difficulty writing and classifying Evaluation and Synthesis items during the time the multiple-choice format was in effect. The Synthesis process essentially involves the re-combination of elements into a new structure which provides a unique communication. Our attempts to cast this process into the forced-choice format generally led to items for which the alternatives were such recombinations and the sought for response was the identification of the one which appeared best or the one which could have been created solely on the information which appeared in the passage. Typically these items were classified as Analysis, but sometimes as Evaluation. The Analysis classification was justified on the basis that the principles in reading passage had to be identified and then the alternatives to the items had to be analyzed to determine which of them contained these principles but in a new form. The Evaluation classification was sometimes given on the basis that the best one of the alternatives could be selected only on the basis of invoking some external or internal set of criteria by which to judge the relative goodness of each of them.

The multiple-choice items which were intended to assess Evaluation were most frequently rated as Analysis. Usually the items contained alternatives which embodied the product of the Evaluation process and the student was requested to choose the best one of them. Therefore the raters regarded the required process as Analysis because the criteria for judgment were generally included in the stem; consequently, there was no need for the student to furnish the criteria which seems to be an essential part of Evaluation. Another difficulty which we encountered with Evaluation items was our failure to realize that by-and-large the criteria for Evaluation must be supplied from a source other than the reading passage; e.g., evaluation according to an external set of standards. We seemed to ignore this, probably because we believed that the passage should contain all relevant information except for the process itself.

For an extended period the project staff developed ad hoc principles for determining when an item should be classified as Evaluation or Synthesis. Each principle generally had a life dating from the time the agreement was reached until the next time it was considered. It is for these reasons that the forced-choice format was abandoned for these levels. This decision came quite late and as a result the Evaluation and Synthesis items which appear in the four test booklets are undoubtedly not as good as they should be nor as good as they could have been had more time been available for trying different techniques of writing such items.

E. Administration of Preliminary Forms

The discussion of the administration of preliminary test forms is limited to the four forms which were finally adopted.

Each form was administered at least three times to groups of secondary school students. In every administration, tenth- and eleventh-grade students were tested, but in about only half of the administrations were ninth- and twelfth-grade students included. This limitation was due partly to our belief that we could obtain the needed data from the mid-groups, the fact that it was difficult to obtain twelfth-grade students because school officials felt that too much of their time had been devoted already to a required state testing program and several national and special testing programs, and the fact that it was awkward to obtain many ninth-graders because that grade level is in neighborhood and feeder schools which are generally small.

All tests were administered by the project staff. Two members were used for a thirty-student group and one additional staff member was used for each additional thirty students. Local teachers and guidance counselors were sometimes used as proctors, but only when there was a forewarning that their presence might be useful in preserving the order of the group.

During each administration, observational records were kept by proctors and, in the majority of instances, students were asked to provide feedback after the tests had been administered.

The proctors entered in their records all student questions about test directions, reading passage, and test items. The questions about items were identified according to the intended level of the item. The frequency of questions was positively related to the level of the item. Proctors noted the time during the test when student inattention seemed to occur. One proctor usually collected information about the apparent incidence with which students referred to the reading passage during the test. When possible, the time was noted at which some students apparently quit the test, i.e., the time at which they "gave up" prior to expiration of the time limit. The proportion who finished within the time limit was also noted.

Students were asked to supply information either by checking items on a questionnaire, giving oral responses in a group session, or writing responses to relatively unstructured questions about the following topics; whether they had previously encountered the content of the reading passage, whether the reading passage seemed to be too difficult for them, whether the technical vocabulary in the reading passage handicapped their reading and responding, whether the items seemed to be too difficult, and whether they believed that previous coursework in the general area of the content would have made the test less difficult and increased their scores, and the ways in which they believed the test differed from those they had taken in the past.

The manner in which these two sets of information and their test responses were used in the revision of tests is described in the next section.

F. Analysis of Preliminary Forms

The two major sources of information used in the revision of test forms were (a) observations of students and their reactions to the tests, and (b) statistical analyses

of their responses to the tests. To a minor extent information gained through recordings of student oral problem solving processes were used but this information proved not to be very helpful. Although decisions to revise the tests were usually based on information from both sources, each is presented separately because it enables identifying more pointedly some problems which arise when constructing taxonomy-type tests.

Proctor Observations and Student Reactions

Many changes in the test directions were made because of proctor observations and student reactions. The directions were altered to emphasize that students could refer to the passage anytime during the test administration. In some versions this instruction was written on to each page of the test booklet. Even when students were so instructed they were suspicious, apparently wondering whether it was fair or whether the test could be any good if one was permitted to look up the answers.

The directions were altered to inform students pointedly that, although they could refer to the passage and find some of the answers, the answers to some of the items could be reached only by careful consideration of the content and that no specific answers to these appeared in the passage. The need for this direction came from student reactions that the passages were incomplete because some of the answers did not appear in them and from proctors who observed students searching endlessly for specific answers which were not there.

The directions for the free-response Evaluation and Synthesis items underwent several revisions to emphasize unmistakably that answers did not appear in the passage and that several different but equally correct answers might be given to an item. Giving divergent responses to items was probably unfamiliar to the students and they were uncomfortable when asked to do so.

The directions were altered to advise students simultaneously that all answers did not appear in reading passage and that they should not spend an undue amount of time on any item. Proctors observed that Synthesis and Evaluation items often received inordinately extended consideration from students. Proctors also believed that students did not attend to the tables, charts, etc., which were in the reading passages, so the directions were written to contain reminders that important information appeared in them.

The reading passages were altered on the basis of student reactions. Chiefly, these changes involved diluting

the technical terminology, creating more but shorter paragraphs and sometimes increasing the readability of sections, and, in one instance, preparing and adding a table so that the content might be more readily understood.

In addition to the difficulties which project staff had in preparing and classifying Evaluation and Synthesis items, students were perplexed by them, spent an inordinate amount of time on each of them, and proctors were more frequently questioned about them than about any other level of item. These factors, along with statistical data which are presented subsequently, caused the staff to abandon the specifications that the multiple-choice format be used for Evaluation and Synthesis and that the total test be brief enough to be administered in one class period.

Also, student reactions of disinterest to one passage caused us to abandon it although it was otherwise well-suited to the purpose and the item analysis data were satisfactory.

Statistical Data

After each preliminary version was administered certain routine statistical analyses and descriptions were prepared for use in revising the form. These included: frequency of response to all alternatives and skips and omits for each item; subtest means and standard deviations; and item-subtest correlations. These were used as follows: items which had extremely low or negative correlations with overall subtest scores were re-written or discarded; alternatives which received no responses were either discarded or re-written to be more plausible; and items for which distractors received a higher response than did the responses were checked carefully, along with the reading passage, to be certain they were not technically inadequate.

The reader will surely be aware that some standard methods of item analysis and item selection were not mentioned in the above description although the necessary data were available. The staff attempted to confine its revisions to those items which were obviously bad by any standard and to be certain that other items were free of defects which could be altered by simply editing them. The reasons for this are presented in the next section.

G. Problems in the Analysis of Item Data

A major problem confronted the staff when item analysis data from the preliminary forms were to be evaluated. The problem was fundamentally theoretical and methodological, and it had distinct ethical overtones. The latter is treated first, then the former.

The ethical problem was simply this: possession of the item analysis data provided the staff with information which bore directly, through a chain of fixed activities, on the ultimate outcome of testing the hypothesis about the structure of the taxonomy. How the analyses were used could finally shape the result of testing the hypothesis. An example will clarify this matter. It is claimed that the taxonomy is hierarchical; consequently the investigators arranged several tests of this hypothesis, one of which was that mean difficulty on subtests would increase as the level of the subtest increased in the structure. Thus, item analysis data provided the opportunity to select items which would determine whether the hypothesis would eventually be accepted or rejected. The staff extricated itself by deliberately not making certain kinds of decisions which are ordinarily made in test construction. For example, items were not selected for inclusion in the subtests on the basis of preconceptions about the desirable range of item difficulties, standard deviations, or mean difficulties. Generally test constructers are urged to achieve a .50 difficulty level in each subtest so that the maximum number of discriminations will be made by items. Adhering to this suggestion would have resulted in a test in which all subtests had equal difficulty and one consequence would be that the hypothesis about the hierarchical structure of the taxonomy would have been rejected. We adopted this working principle: if an item correlated positively with its subtest score and if it was free from technical deficiencies, then it was included in the test; however, it might have been discarded randomly later in order to have the same number of items in each of the subtests for the first four levels.

The theoretical and methodological problems were numerous and, unfortunately, most of them remain unresolved. Our only contribution in this regard is that we isolated them.

First, we did not know how to regard answer sheets on which Knowledge items were answered incorrectly or

omitted. If the relevant information appeared in the passage and if the Knowledge item was answered incorrectly, then numerous questions arose. (Of course, we are omitting from this discussion the fact that students could have overlooked items, entered the wrong response through carelessness, etc.) Two of the prominent ones are: (a) was the stimulus value of the item so low that the student did not recognize that the bit of knowledge which he stored was relevant for responding to this item? and (b) would the student's answer sheet have any value whatsoever on account of this incorrect or omitted response. With regard to the former, we simply assumed that the faulty or absent response was due to our failure to write an item of high stimulus value; i.e., we assumed the fault was in the item and rejected the possibility that the student did not possess the relevant content. So we altered the item in a way that we hoped would make the correct response more obvious than it was. That we were not firm in our assumptions is borne out by two other steps we took. We lowered the reading difficulty of the passages to maximize the student's chance of storing the information in the first place. We placed all Knowledge items at the beginning of the tests and ordered them according to their order of appearance in the passage. The second question--could we use the student's answer sheet if he missed a Knowledge item--was one that we resolved arbitrarily by stating that we could. We regard the decision as arbitrary but our reasoning was that the reading passages were clear and within his grasp, consequently, the fault was due to a poorly written item. Therefore, it would be appropriate to use his responses in the analysis.

All of this might seem rather trivial but the problem is fundamental. Because the taxonomy is reputedly hierarchical and cumulative, then performance at one level is contingent upon performance at the preceding level. In other words, if a student cannot recall the appropriate content; i.e., if he fails a Knowledge item, then the same deficiency should appear in Comprehension items, in Application items, etc. If such answer sheets are accepted for inclusion in the item analysis, then one knows in advance that item difficulties will be increased probably ascendingly through the upper levels. Also, the student's level scores are likely to be misinterpreted because they do not accurately portray his mastery of the several processes. His lack of "Knowledge" prevented him from applying the higher level processes.

So the second major problem was how to regard the item analysis data for each of the levels once it had been obtained. In the typical situation one strives for an average difficulty level of .50 unless he is purposely peaking the test and one also strives to reject items having sharp splits because they do not yield many discriminations and they reduce external validity. But these guidelines do not hold for taxonomy-type tests where one knows on theoretical grounds that difficulties will increase as the level of the item increases and that item splits will depart rapidly from the prescribed 50/50 split. Thus, to use established principles will yield a test which is at variance with the notion of a sequence of contingent tests each of which is more difficult than the next. Guidelines for constructing taxonomy-type tests have not been developed. New principles are needed that would suggest the probable average difficulty of Comprehension items on the basis of known difficulties of Knowledge items, etc., through the taxonomy; that would forecast the number of sharp and nondiscriminating item splits which one might expect on the basis of predicted overall difficulty of a subtest, etc.

These problems became even more complex when attempting to judge the item analysis data from the free response Evaluation and Synthesis items. There was no keyed response for them in the usual sense, they were the most difficult of all items and it would be expected that nearly every item would depart markedly from a 50/50 split <u>and still be valid</u>, and that the responses were scored on a quality scale. We simply had no tools for coping with the data. If the responses were disappointing, we attempted to sharpen the item, but we held no expectancy about the nature of a quality point distribution of a good set of responses to a valid item. We attempted to compare performance over grades for each of these items and tended to select items on which success increased as grade level increased. Whether the increased performance was due to increased mastery of the processes or simply greater maturity is an open question.

H. Preparation of Final Forms

Four final forms of tests were prepared in order to collect data with which to study the hypotheses. Two of these dealt with social science materials--the Lisbon Earthquake and Stages of Economic Growth--and the other two dealt with science materials--Atomic Structure and

Glaciers. The differences between the science and social content are not nearly as sharp as we had planned them to be. Also, they should not be regarded as interchangeable.

Each form consists of two parts. Part A contains directions, the reading passage, and items which reflect the first four levels of the Taxonomy--Knowledge, Comprehension, Application, and Analysis. All Knowledge items appear first, then items for the other three levels appear in random order. There are twenty items for each of the four levels. Administration time is one hour. Part B contains directions and the items for Synthesis and Evaluation. Part A must be furnished to students when they take Part B in order that the reading passage is available to them. Part B contains five Synthesis items and ten Evaluation items. The administration time is one hour.

No studies were conducted on differential performance on Parts A and B with regard to order of administration. Every student in the validation sample took Part A then Part B. Until experiments are done on order of presentation, the staff strongly recommends that they be administered in that order, and that Part B should not be administered unless Part A is too.

The method of selecting items for each level for each form was as follows: On the basis of item analysis data from preliminary forms, students' comments, and proctor observations all items were ridded, hopefully, of technical deficiencies and items were discarded which failed to correlate positively with the subtest of which they were a part. When these two steps were achieved, then the item was placed in a pool of items for that subtest and form. When the number of items to appear in the final form was determined, items were sampled from the pool until the needed number was obtained. It should be noted that the pools of items became increasingly smaller as the level of items increased.

In the case of Synthesis and Evaluation items there was not a great deal of choice because of the limited number of items we were able to produce and because we did not know how to interpret with any confidence the item analysis data which we obtained, except for the item-test correlations. However, each item was answered correctly more frequently as the grade level of the respondents increased.

IV. TEST ADMINISTRATION AND SCORING

A. Test Administration

This section describes the selection of schools, testing schedules for the taxonomy-type tests and the cognitive aptitude tests, the format of the testing materials, test administration teams, test directions, student reactions to the testing, and the procedures for scoring the tests.

All taxonomy-type tests and the cognitive reference tests were administered in April and May of 1965.

Selection of Schools

Data from the Florida State-Wide Ninth-Grade Testing Program (which is compulsory for Florida public school students) were used to select schools for the taxonomy testing program. Schools were selected according to these criteria: (a) they were located in Florida; (b) they were willing to participate; (c) the school mean scores achieved by their students fell between the 40th and 60th percentiles on verbal and quantitative ability as determined by state norms (but this had to be breached as described subsequently); and (d) their pledge of suitable test administration conditions and cooperation by local school personnel.

Only Florida schools were included in the sample for several reasons. First, a sample of students from the state could be drawn which would have characteristics similar to a national sample. Second, scores on the battery of achievement and ability tests from the two state-wide testing programs were available only for Florida students and these tests results would be useful in postproject studies of the concurrent validities of the taxonomy-type tests.

Criterion (c) was established because it would yield schools whose students display a wide range of ability. After records were searched to identify schools which met the criterion, letters were prepared and sent to county school officials to request their participation. The letter included a summary of the project and information about the number of students and the time from them which were needed. Only one

invitee could fit the testing program into his school calendar. Therefore the criterion was expanded to include a wider score range and a new set of invitations was extended. Eventually enough invitations were accepted to yield the number of students needed. However, the ability level represented by these students is above average. The students from two of the schools are markedly above average. An explanation regarding these two schools is presented later.

Table 7 contains a list of schools which were in the sample, their locations, and the number of taxonomy forms which were administered in each.

TABLE 7

Schools in the Taxonomy Testing Program

School	City	County	Number Taxonomy Adm. Each Subject
Cocoa Beach High	Cocoa Beach	Brevard	4
Melbourne High	Melbourne	Brevard	4
Rickards Jr. High	Ft. Lauderdale	Broward	4
North East High	Ft. Lauderdale	Broward	4
Walker Jr. High	Bradenton	Manatee	2
Manatee High	Bradenton	Manatee	2
South West Jr. High	Lakeland	Broward	4
Winter Haven High	Winter Haven	Polk	2
*Mainland Jr. High	Daytona Beach	Volusia	4
*Mainland Sr. High	Daytona Beach	Volusia	4

*Also designated for the Kit of Reference testing.

Table 8 contains the mean percentiles on the School and College Ability Tests and the Metropolitan Reading Test of the ninth-grade students in schools which were in the sample and in feeder schools which supplied students to schools which were included in the sample. Current test data were available only for ninth-grade students; consequently any judgment about the mean percentile scores attained by tenth- through twelfth-grade students in a sample school must be extrapolated from the current ninth-grade means of students in that school or, if the school does not contain a ninth grade, from the principle feeder school. The staff believes the ninth-grade means represent the overall school means because no between school differential dropout factors were noted.

TABLE 8

Aptitude and Achievement Means for Sample Schools Having Ninth-Grades and Feeder Schools for Sample Schools Without Ninth Grades

School Number	Sample School?	Feeder School?	SCAT Means V	Q	T	Metropolitan Reading Test
1	Yes	Yes	58	64	59	48
2	No	Yes	42	47	48	48
3	Yes	Yes	58	75	69	62
4	Yes	Yes	58	47	59	62
5	No	Yes	69	47	65	48
6	Yes	No	91	86	89	90
7	Yes	Yes	91	93	92	90
8	No	Yes	58	34	65	48

It was stated earlier that two schools in the sample had high mean scores on the School and College Ability Test. One of these, a junior high school, did not feed to a senior high school in the sample; therefore, this high ability characteristic should not be generalized to any tenth-, eleventh-, or twelfth-grade group. The other high ability school was also a junior high school and was a feeder school for a sample school. However, the overall effect of it as a feeder school is somewhat offset by a second feeder school which had substantially lower scores. Both were feeder schools for the same senior high school in the sample. These schools were included because they were needed and because they met all criteria except the one about ability. However, the selections were made only after winning the assurance by principals that they would select students to participate who would represent as complete a range of ability as possible. The number of students completing each form in these schools was relatively small and an examination of the test performances indicates that no bias was probably introduced by their inclusion. In the majority of schools, all students were scheduled to take all four taxonomy forms; but in a few instances students were scheduled to take only two forms. Therefore, scores were available for a minimum of two forms and for four forms for each student.

The enrollment of the sample schools ranged from approximately 200 per grade to slightly more than 500 per grade. The smallest group completing any form was forty and the largest group was slightly more than 500. It was the opinion of the project staff that the socio-economic status of these students was predominantly middle class. Very few students of low socio-economic status were included because of the manner in which schools were selected.

Testing Schedules

The order for administering the taxonomy forms was determined within these restrictions: (a) the first form to be administered would not be Atomic Structure, (b) test content; i.e., social science and science, was alternated over administrations; and (c) within restrictions (a) and (b), the order was established randomly.

In every administration, Part A (Knowledge, Comprehension, Application, and Analysis) of a given taxonomy-type test was followed immediately by administering

Part B (Synthesis and Evaluation) of that form. The students had a short break between Parts A and B. In most cases, the two parts were administered in successive class periods. In no instance were two forms administered without a substantial intervening period of time. The minimum break was a lunch period, and in most cases the forms were administered on successive days.

Administration of the KIT Tests

In one county the Kit of Reference Tests (KIT) was administered to students in grades nine through twelve in addition to the taxonomy-type tests. The KIT tests are designed to measure specific cognitive factors.

The taxonomy and KIT testing schedule was designed to (a) maximize the time between the administration of taxonomy forms; (b) maximize the time between the administration of pairs of KIT subtests which measure the same factor; (c) minimize the total testing time required of each pupil; and (d) fit the testing schedules within the time periods which were made available by the schools. Testing of ninth-grade students was completed within a period of five days, with alternate morning and afternoon testing periods. For students from grades ten, eleven, and twelve, testing was scheduled only for mornings and was distributed over a three week period.

Format of Testing Materials

Each taxonomy form consisted of two parts: Part A contains the reading passage and 80 multiple choice items evenly divided among Knowledge, Comprehension, Application, and Analysis. Part B contains five free-response items for Synthesis and ten free-response items for Evaluation. Directions are printed on the cover page of both parts. The tests were produced by the photo-offset duplicating process. The completed forms measured seven inches by ten inches. The test forms in their entirety are reproduced in Appendices A through D.

Test Administration Teams

All tests were administered by teams consisting of project staff members and local school personnel. One member of the project staff was responsible for the

actual test administration and the other members served as proctors. Observation of behavior during test administration was hampered in four schools where students remained in classrooms and received directions via public address systems. In all four schools, this method of testing was common and created no unusual problems.

In all schools, meetings were held with the faculty to discuss the project, explain its goals, and to answer any questions about the project or testing procedures. In general, all local school personnel were most cooperative. A negative attitude toward the testing program on the part of teachers in one school was detected near the end of the testing program. There is a possibility that some of this negative feeling was reflected by the pupils in their performance.

Directions to Students

The actual administration of the tests followed generally accepted procedures. Specific directions which were read to the students and other details of administration appear in Appendix E.

The KIT battery consisted of 37 subtests, which measured seventeen specific factors. A minimum of two tests per factor, was used. These subtests were grouped into batteries on the basis of the amount of time available for administration and to maximize the variety of factors at each sitting. These batteries were mimeographed with the exception of subtest R-3 (Ship Destination) and subtest Rs-2 (Logical Reasoning) which were obtained from the publisher.

Student Reactions

The students spent the first ten to fifteen minutes reading the passage. The directions did not indicate a specific time limit for reading the passage. The instruction was to read the passage and then begin to answer the items. Few students attempted to answer items before reading the entire passage. Questions which students asked while reading the passage dealt mainly with vocabulary. Most of the questions about the reading passage were asked while the students were responding to the items of Part A. The majority of these questions dealt with the charts and graphs. (The students found the Table of Atomic Structure to be quite difficult to understand.)

In general, the students reacted to Part A in the same manner as one would expect for conventional tests. The only major problem seemed to be their inability and/or unwillingness to interpret the charts and graphs. Other problems arose due to inadequate vocabularies.

The most frequent comment occurring during Part B was "I don't know the right answer" which was frequently given in response to requests that they try to complete all items. The proctors were frequently asked to read responses and react to them, but this was not done.

Discipline was not a major problem. The presence of local school personnel probably accounts for it. Students who completed the test before time expired were asked to read other materials which they had been requested to bring with them. A few isolated instances of random marking were observed and those students were requested to try to answer the questions or to leave the answer positions blank.

Whenever possible, time was allowed for students to ask questions after completing the last taxonomy form. Their questions might be categorized as follows: Requests for correct answers to specific items of Part A; requests for opinions about specific responses given to Synthesis items; and the purpose of the tests.

The examiners overheard several conversations by students about the tests. Generally, the topic was the Synthesis items. High ability students, as identified by their teachers, seemed to regard the taxonomy-type tests as a challenge and appeared to enjoy the Synthesis and Evaluation items more than they did the others.

The motivational level of the students was judged to be at a slightly lower level than is desirable. In all schools, students asked the examiners if the results would be placed on their permanent records. The students were told that the results would be returned to their schools and the principals would decide what use to make of them. The fact that these tests were experimental tests was indicated in the introductory comments and appeared on the cover sheet of the booklets. Many students drew the conclusion that the results of the tests would, therefore, be meaningless to them.

B. Test Scoring

The multiple-choice items were scored on an IBM 1230 scoring machine. The free-response items were scored by persons who were hired specifically for that purpose and who worked directly under the supervision of project staff members. This scoring staff was selected according to the following criteria: no prior experience scoring free-response items, completed at least three years of college, and attained high scores on a pre-employment test which was one of the taxonomy forms.

Their initial training consisted of taking a taxonomy form as a pre-employment test, studying Appendix A of the Taxonomy, and reviewing scoring criteria developed by the project staff.

The Synthesis items were scored on a scale of 0-4 and Evaluation items were scored on 0, 1, 2 scale. A comprehensive discussion of scoring Synthesis and Evaluation items appears in Appendix E.

Scoring criteria in the form of sample responses to Synthesis items for one form were developed by project staff and presented to the scoring staff. They were instructed to read the criterion samples for a given item and tentatively to assign scores to a limited sample of responses. A project staff member reviewed their scoring. When he felt that the scoring was accurate they were instructed to assign scores. In the early stages, each scorer worked on a different item of the same form. This procedure wasted training time because each scorer had to be trained separately. Also, it created additional problems in quality control because the project staff members had to review responses to several different items which had been scored simultaneously. Thus, it was judged that there would be better control if all scorers worked on the same item simultaneously.

For these reasons the procedures were modified for subsequent taxonomy forms. The project members selected tentative criteria for each question and presented them to the staff for discussion. During this meeting, each scorer reviewed several papers and presented questionable responses to the remainder of the staff for discussion and decision. A set of sample responses was developed by this process and it was used as criteria.

The procedures followed for scoring Evaluation items were essentially the same as those which were developed for scoring Synthesis items.

Additional quality control was maintained in an informal way through staff discussion of responses which were difficult to score. These responses were presented to the project staff members for decision. Also, the project staff rescored answer sheets which had already been scored by the scoring staff. A conservative estimate of the number of scorable responses which were double-scored would be one item in ten for Synthesis and one in twenty for Evaluation.

The KIT tests were scored by a trained staff different from the taxonomy scoring staff. This staff had been trained earlier when they were employed on another project. Quality control was maintained by randomly selecting papers and having project staff rescore them.

Complete information about the scoring of the taxonomy-type tests including sample responses appears in Appendix E.

V. STATISTICAL DATA ABOUT THE TAXONOMY FORMS

The purpose of this section is to present summary data about the taxonomy-type tests. These data, in conjunction with certain other data which are presented in the Appendices, are intended to give to potential users of the test forms as complete a statistical representation of the tests as is possible.

The means and standard deviations by grade and taxonomy level over schools, some reliability data, and the intercorrelations of levels by grade and over forms and schools are presented here, along with some cautions in their interpretation.

Appendices A through D, each of which is devoted exclusively to one of the test forms, contain for each form the item distribution data, item-level correlations, percentile distributions, and reliability estimates for levels Knowledge through Analysis. Reliability data for Synthesis and Evaluation are contained in Appendix E.

Most of the information presented in this section serves as the basis for testing the hypotheses about the hierarchical cumulativeness of the taxonomy and the transcendence of the processes, the results of which appear in Section VI.

A. Means and Standard Deviations

The test forms were administered in ten different schools to a total of twenty-three different grade groups. The smallest grade-group consisted of forty-one students and the largest consisted of 508 students.

The means and standard deviations by level, form, and grade over schools appear in Table 9. The lowest and highest school means for each level by grade appear in Table 10.

The chance mean for each of the first four levels is five (twenty four-choice items for each level) and the chance means on Synthesis and Evaluation are indeterminate because both sets of items are in a free-response format.

The tabular data reveal that means increase with grade level, means decrease as level increases within grade, and that there is no clear single pattern of change of standard

TABLE 9

Means and Standard Deviations for
Taxonomy Norms within Grade and Over Schools

	Atomic Structure		Glaciers		Lisbon Earthquake		Economic Growth	
			Grade 9					
K	12.98	4.08	15.66	3.71	15.87	3.28	14.32	4.18
C	6.94	2.90	7.71	3.24	10.39	3.60	8.19	3.53
Ap	6.38	2.88	8.48	3.71	8.54	3.26	8.13	3.51
An	5.22	2.56	6.27	2.79	6.78	3.08	6.24	2.78
Sy	0.49	0.90	3.06	2.03	2.43	1.78	1.23	1.29
Ev	1.20	1.44	4.17	2.90	1.93	1.98	0.75	1.11
	N = 1184		1223		1199		1199	
			Grade 10					
K	14.05	4.24	16.42	3.62	16.86	2.98	15.11	4.13
C	8.21	3.37	8.86	3.42	11.65	3.87	9.40	3.86
Ap	7.83	3.39	9.55	3.83	9.93	3.68	9.23	3.63
An	6.35	2.95	7.25	2.96	7.98	3.53	7.06	3.09
Sy	0.94	1.41	3.81	2.23	3.60	2.44	1.79	1.84
Ev	1.56	1.78	4.64	3.23	2.25	2.38	.97	1.30
	N = 1422		1436		1455		1384	
			Grade 11					
K	13.95	4.66	17.06	3.26	17.33	2.65	15.91	3.56
C	8.46	3.90	9.65	3.52	12.47	3.86	10.18	4.18
Ap	8.21	3.98	10.58	3.90	10.58	3.62	10.02	3.93
An	6.82	3.45	7.61	3.01	8.64	3.63	7.69	3.39
Sy	1.17	1.54	3.98	2.40	3.87	2.49	2.39	2.02
Ev	1.69	1.95	5.16	3.31	2.84	2.55	1.18	1.41
	N = 1261		1309		1303		1291	
			Grade 12					
K	15.26	3.63	17.41	2.91	17.78	2.24	16.21	3.55
C	8.92	3.55	9.91	3.59	13.37	3.49	10.70	3.93
Ap	8.71	3.66	10.66	3.88	11.45	3.42	10.56	3.75
An	7.02	3.35	7.90	2.95	9.50	3.60	8.14	3.40
Sy	1.40	1.60	4.04	2.34	4.43	2.40	2.40	1.86
Ev	2.10	2.06	5.57	3.16	3.03	2.54	1.38	1.49
	N = 1197		1186		1192		1194	

TABLE 10

Range of School Mean Scores within Grade

	Atomic Structure		Glaciers		Lisbon Earthquake		Economic Growth	
				Grade 9				
K	12.04	14.12	15.14	16.60	15.41	16.71	12.70	15.44
C	4.85	7.69	6.63	8.39	9.41	11.28	7.93	9.13
Ap	4.50	7.04	7.32	8.86	8.31	9.41	7.77	9.15
An	3.25	5.69	5.11	7.40	6.03	7.48	5.53	6.54
Sy	.30	.70	2.22	3.75	2.01	2.97	1.03	1.60
Ev	1.10	1.22	3.11	5.45	1.31	2.56	.36	.93
				Grade 10				
K	13.16	16.10	15.08	17.16	15.68	18.42	13.49	15.98
C	5.13	11.02	6.98	10.14	9.04	14.96	7.85	10.87
Ap	5.16	10.02	7.76	11.57	8.10	13.10	8.34	10.70
An	3.67	9.08	5.50	8.16	6.42	11.34	5.02	8.70
Sy	.55	1.49	3.22	4.56	2.22	7.24	1.22	2.43
Ev	1.33	2.72	3.57	7.23	1.31	5.74	.48	2.17
				Grade 11				
K	11.13	16.53	16.52	17.53	16.75	17.72	14.21	16.68
C	6.03	11.55	8.75	10.67	10.94	13.46	7.26	11.48
Ap	5.85	11.32	9.57	11.53	9.28	11.81	7.46	12.32
An	4.94	9.11	7.15	8.20	7.52	9.63	5.44	8.75
Sy	.64	1.90	2.87	4.80	2.57	5.96	1.42	3.17
Ev	1.35	2.86	4.47	7.11	1.86	5.53	.65	2.00
				Grade 12				
K	13.86	15.93	16.75	17.94	17.20	18.12	13.93	17.01
C	6.02	10.31	9.03	10.63	12.40	13.81	9.11	11.53
Ap	6.08	9.61	9.57	11.35	10.68	11.84	9.06	11.23
An	4.27	1.85	7.36	8.46	8.53	9.74	6.60	8.70
Sy	.54	1.67	3.04	4.92	2.95	5.98	1.58	3.10
Ev	1.36	3.20	4.88	6.94	1.64	5.02	.82	1.85

deviations over grades or over levels within grades. These general tendencies are to be expected from the joint standpoint of the imputed structure of the Taxonomy and the nature of the tests which were constructed.

The exception to these tendencies are the reversal from expectancy of the Comprehension and Application means at all grade levels on the Glaciers test and the general lack of order of the Synthesis and Evaluation means. The former is due to some Comprehension items which dealt with obscure information in the reading passage. The latter can more readily be commented about than explained. All Synthesis items preceded the Evaluation items, so there was an element of practice which might have augmented the responses to the Evaluation items. The Evaluation items were more structured and could in some instances be attempted on a common sense basis. The Evaluation items called for the evaluation of situations on the basis of internal and external criteria, the latter being external to the reading passage and therefore not amenable to control by the test authors. The Synthesis items were unique, according to student testimony, whereas the Evaluation items were regarded as more commonly encountered in school settings than were the Synthesis items.

Reliability data for the taxonomy-type tests appear in Table 11. The coefficients are computed over grades. Within grade coefficients are probably substantially smaller than those reported for Knowledge through Analysis. The Synthesis and Evaluation coefficients within grade are probably of the same magnitudes as those which are presented.

C. Intercorrelations

Matrices of intercorrelations were computed for the complete battery of taxonomy-type tests. A total of twenty-four scores (six levels by four forms) were available for students who took all tests; however due to student absences and our scheduling some students to take only two forms, a substantial number of the test records were incomplete. Consequently, it was necessary to use a missing-point correlation routine for developing the matrices. A 24 x 24 matrix was computed on students from each of grades nine through twelve. These matrices are presented in Tables 12, ninth grade; 13, tenth grade; 14, eleventh grade; 15, twelfth grade; and 16, all grades combined.

The number of cases contributing to each correlation does not appear in these tables, but the minimum number of students contributing to any coefficient is included. This

TABLE 11

Reliability Over Grades of Taxonomy Scores*

	Atomic Structure		Glaciers		Lisbon Earthquake		Economic Growth	
	Raw Score	Corrected Score	Raw Score	Corrected Score	Raw Score	Corrected Score	Raw Score	Corrected Score
K	.836	.929	.816	.919	.758	.885	.824	.924
C	.693	.821	.677	.815	.743	.862	.745	.867
Ap	.694	.826	.731	.857	.689	.834	.731	.862
An	.632	.760	.539	.706	.684	.824	.614	.763
Sy	.89		.72		.71		.79	
Ev	.83		.75		.81		.72	

*For K, C, Ap, and An, KR 20's are reported and each is based on approximately 5,000 students. For S and E, the coefficients are interjudge reliabilities and the method of computation is described in Appendix E.

TABLE 12

Intercorrelation Matrix

All Forms - Grade 9

	Atomic Structure						Glaciers						Lisbon Earthquake						Stages of Economic Growth					
	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev
K		372	322	120	329	363	594	438	426	251	242	290	559	487	493	406	417	343	600	478	476	295	317	287
C			556	485	169	268	316	386	377	316	209	280	307	397	382	316	335	263	226	373	414	342	213	216
Ap				443	214	278	324	359	404	278	205	258	300	387	376	324	334	227	257	386	393	333	218	224
An					064	121	131	265	250	274	110	154	106	208	224	192	192	098	056	272	254	246	109	156
Sy						375	293	281	297	173	224	239	294	290	337	283	375	300	282	354	329	236	335	244
Ev							347	289	314	154	182	380	322	506	357	304	429	399	380	362	377	258	312	334
K								380	437	227	275	391	590	471	443	390	438	330	622	472	479	301	361	307
C									631	589	312	286	373	510	498	405	373	220	339	538	523	460	339	320
Ap										578	387	343	426	531	570	450	412	264	385	582	577	493	336	371
An											289	197	277	396	417	323	262	103	161	386	350	329	193	207
Sy												234	241	307	304	226	256	114	189	299	326	246	233	216
Ev													313	338	293	284	334	352	358	392	374	263	301	344
K														544	517	419	481	318	467	419	417	295	316	266
C															642	573	475	304	420	538	520	440	332	319
Ap																565	499	315	414	518	513	394	341	301
An																	441	273	348	496	441	421	354	306
Sy																		372	379	450	442	338	414	388
Ev																			347	334	321	266	325	331
K																				496	482	246	368	319
C																					698	578	426	414
Ap																						522	444	421
An																							304	272
Sy																								318
Ev																								

A missing point correlation program was used to compute these intercorrelations. The number of subjects contributing to each value varies, but in all cases the number was greater than 750.

TABLE 13

Intercorrelation Matrix

All Forms - Grade 10

	Atomic Structure						Glaciers						Lisbon Earthquake						Stages of Economic Growth					
	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev
K		528	483	364	401	374	610	477	484	355	285	412	519	468	511	466	458	330	602	494	482	329	380	342
C			660	504	401	362	409	520	484	390	317	397	373	498	489	485	449	348	389	518	492	448	376	369
Ap				571	441	346	427	510	528	443	349	399	347	470	463	462	431	320	359	514	476	438	407	372
An					367	300	305	465	427	320	290	318	275	426	398	405	420	314	283	468	398	389	368	334
Sy						449	327	384	374	267	375	391	306	372	364	375	502	347	322	396	368	286	444	323
Ev							343	358	344	228	269	500	297	325	379	378	392	380	533	346	336	276	383	352
K								496	507	361	306	437	640	502	509	440	468	321	630	477	483	332	369	334
C									689	574	386	392	460	588	589	539	489	338	414	509	560	501	454	393
Ap										593	408	406	452	568	598	540	463	295	401	585	558	508	422	410
An											326	284	323	428	447	395	329	213	262	493	466	421	320	245
Sy												296	302	360	386	344	418	291	209	387	352	329	371	277
Ev													384	385	378	381	434	411	389	422	414	310	427	415
K														565	548	485	482	341	482	431	433	310	315	327
C															721	671	557	393	411	561	525	503	403	382
Ap																684	538	356	435	593	579	494	428	410
An																	550	381	376	560	544	471	440	408
Sy																		439	433	498	476	418	518	431
Ev																			299	347	329	266	326	330
K																				532	508	349	370	361
C																					733	649	490	455
Ap																						580	443	420
An																							390	378
Sy																								397
Ev																								

A missing point correlation program was used to compute these intercorrelations. The number of subjects contributing to each value varies, but in all cases the number was greater than 1100.

TABLE 14

Intercorrelation Matrix

All Forms - Grade 11

	Atomic Structure					
	K	C	Ap	An	Sy	Ev
K		581	566	440	421	380
C			775	693	428	368
Ap				695	454	336
An					378	322
Sy						389
Ev						

	Glaciers					
	K	C	Ap	An	Sy	Ev
K	561	478	484	347	298	381
C	365	545	494	422	267	389
Ap	391	566	539	442	286	375
An	303	476	445	381	244	328
Sy	300	374	383	285	370	364
Ev	273	280	253	191	156	481
K		506	532	353	278	373
C			690	600	352	366
Ap				607	389	385
An					287	297
Sy						223
Ev						

	Lisbon Earthquake					
	K	C	Ap	An	Sy	Ev
K	500	473	474	433	472	418
C	413	502	496	480	491	454
Ap	398	516	502	513	486	428
An	325	446	441	456	422	401
Sy	297	383	426	445	566	434
Ev	282	290	304	299	359	439
K	517	447	435	380	387	313
C	403	572	551	506	467	384
Ap	420	573	592	534	473	358
An	302	461	475	433	349	271
Sy	242	334	354	320	377	208
Ev	327	354	365	327	389	481
K		567	550	494	444	316
C			717	678	538	384
Ap				704	563	390
An					542	381
Sy						487
Ev						

	Stages of Economic Growth					
	K	C	Ap	An	Sy	Ev
K	568	493	504	332	373	368
C	427	609	579	511	415	424
Ap	395	620	577	525	413	426
An	338	549	496	468	352	361
Sy	324	416	398	329	456	386
Ev	317	308	331	205	334	355
K	536	458	488	317	354	332
C	366	619	608	506	434	402
Ap	370	617	615	504	418	417
An	220	501	483	424	293	335
Sy	154	282	298	223	286	247
Ev	364	382	379	280	350	412
K	541	473	487	333	340	322
C	434	620	606	529	459	411
Ap	421	594	605	504	439	398
An	396	575	580	490	449	392
Sy	407	519	514	455	579	442
Ev	333	413	425	360	398	442
K		483	500	307	386	337
C			772	705	517	471
Ap				626	514	463
An					423	404
Sy						412
Ev						

A missing point correlation program was used to compute these intercorrelations. The number of subjects contributing to each value varies, but in all cases the number was greater than 1050.

TABLE 15

Intercorrelation Matrix

All Forms - Grade 12

	Atomic Structure						Glaciers						Lisbon Earthquake						Stages of Economic Growth					
	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev
K		464	491	361	406	361	508	446	441	342	219	376	473	437	421	414	427	355	476	426	445	355	323	357
C			731	683	435	305	354	495	507	386	297	282	344	460	425	416	398	318	264	496	475	460	396	370
Ap				662	457	315	379	574	565	441	313	309	366	496	444	442	405	301	288	543	504	493	411	433
An					393	268	283	471	451	386	293	252	305	400	362	388	355	296	210	470	448	459	375	399
Sy						396	322	401	393	295	308	337	311	360	350	406	495	341	284	365	373	327	419	384
Ev							258	314	322	239	175	414	287	318	333	336	357	395	241	280	273	293	253	377
K								456	508	373	254	561	555	457	462	427	398	297	490	425	449	364	337	321
C									709	588	335	374	440	564	530	538	426	331	332	545	543	493	379	373
Ap										610	381	349	448	565	548	543	447	306	279	547	513	510	353	420
An											332	295	317	448	425	439	352	264	174	406	436	392	287	267
Sy												283	238	307	300	327	353	201	120	278	284	239	313	224
Ev													323	349	321	354	386	449	355	354	349	325	326	390
K														526	544	497	435	300	374	450	447	386	334	329
C															686	661	474	301	261	557	518	527	386	382
Ap																652	489	286	244	512	495	480	350	391
An																	525	327	287	578	518	521	389	410
Sy																		380	317	475	462	408	454	420
Ev																			301	346	326	270	303	371
K																				463	464	341	331	350
C																					739	697	456	459
Ap																						651	468	433
An																							421	405
Sy																								385
Ev																								

A missing point correlation program was used to compute these intercorrelations. The number of subjects contributing to each value varies, but in all cases the number was greater than 950.

TABLE 16

Intercorrelation Matrix

All Forms - All Grades

	Atomic Structure						Glaciers						Lisbon Earthquake						Stages of Economic Growth					
	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev
K		514	496	358	412	386	582	475	474	346	284	389	531	490	501	456	469	379	572	490	494	348	369	358
C			712	648	415	355	381	516	490	404	301	368	381	493	481	463	458	380	353	537	522	475	401	383
Ap				632	450	348	400	540	537	436	324	372	377	501	483	479	460	360	355	557	525	438	419	406
An					371	294	289	462	432	364	276	305	285	416	403	412	401	328	261	488	444	437	368	356
Sy						426	331	400	396	294	361	373	326	395	410	428	534	390	327	417	404	339	459	376
Ev							317	331	323	227	215	467	312	336	362	359	400	420	329	342	347	283	342	377
K								461	513	351	298	409	603	497	491	436	448	334	595	479	495	350	377	340
C									690	608	372	382	445	586	574	533	481	356	391	605	584	516	442	400
Ap										614	411	394	459	583	602	546	483	337	386	603	587	525	418	426
An											331	295	333	461	472	430	365	249	235	477	459	419	314	292
Sy												277	284	356	370	340	392	234	193	338	338	283	334	253
Ev													364	389	378	372	420	450	384	430	401	319	380	411
K														577	564	498	486	340	497	467	472	353	353	333
C															719	675	551	379	417	598	573	530	439	406
Ap																684	562	369	417	588	583	504	436	413
An																	558	376	385	584	558	509	452	417
Sy																		453	414	522	510	446	549	449
Ev																			340	393	384	327	375	397
K																				512	508	337	385	353
C																					753	680	509	472
Ap																						619	497	455
An																							425	397
Sy																								409
Ev																								

A missing point correlation program was used to compute these intercorrelations. The number of subjects contributing to each value varied, but in all cases the number was greater than 3850.

minimum number is 750 students for the by-grade tables and 3850 for the over grades table.

D. Cautions in Interpreting Results

Because the taxonomy is reputedly structured from simple to complex basis and is cumulatively hierarchical the data from tests constructed according to its structure should have rather unconventional characteristics. Furthermore, the four test forms which were constructed have such short subtests in terms of numbers of items that differences between adjacent level scores will be slight. Both of these factors make it mandatory that the resulting data be interpreted cautiously and with a full knowledge of their special characteristics.

Interpreting Means

The tabular data show that the means decrease as level increases within grade and that the means increase within level over grades. It should also be noted that the score distance between adjacent means within grade are generally less than the standard deviation of the scores from either level. It should also be noted that the score difference between means within level from adjacent grades is also generally less than the standard deviation of scores at either grade level.

One consequence of these means and standard deviations, coupled with the reliability data, is that statements of differences must be confined to group differences. The data are simply not sufficiently precise to warrant intraindividual or interindividual comparisons. If such comparisons are attempted, then it should be done only after actually calculating the significance of the differences which is to be interpreted.

Interpreting Correlations

The reader should interpret the intercorrelations with whatever caution or special insight which comes from knowing that all, except a few, bivariate distributions of level scores were triangular. The assumption of a normal bivariate distribution on which the Pearson product-moment correlation rests was not met by the taxonomy data. Nevertheless, product-moment correlations were computed because no other measure of association which would be completely justifiable could be identified. Because the entire routine for testing the hypotheses and the psychological structure of

the taxonomy is based on the analyses of correlational data, whatever caution is used when interpreting these intercorrelations should certainly be carried over and applied to the entire contents of Section VI.

If the structure of the taxonomy is valid, then it follows that an individual's true score at one level will establish the ceiling of his true score at the next higher level; i.e., an individual's score at one level of the taxonomy will be equal to the highest score he can thoretically attain on the next higher level. In short, the bivariate distributions of scores will be triangular.

Norms

There appear in Appendices A through D percentile equivalents of raw scores for each level of the taxonomy forms. These were prepared and included because it is the conventional thing to do. We believe they should be ignored. At best, they merely describe the frequency data in terms which probably are more meaningful to the layman than are the raw scores.

We have strong feelings about the inappropriateness of the percentile equivalent scores for these reasons.

First, the range of raw score points is too narrow to warrant calculating percentiles or, for that matter, any normative score.

Second, the underlying raw score distributions are not normal. If the taxonomy structure is valid, then score distributions for levels should be increasingly skewed as level increases. However, the theoretical distribution for Knowledge scores should either be normal or there should be no distribution whatever. The latter would be ideal for the tests we constructed.

Both the first and second reasons are based on characteristics of the data which lead to tremendous differences between percentile equivalents of adjacent raw scores.

Third, we are unaware of the existence of any norm which might be used meaningfully with taxonomy-type data. An individual's scores on the several taxonomic levels are contingent scores; i.e., his score on Knowledge defines the limit of his score on Comprehension, his score on Comprehension defines the limit of his score on Application, etc. It seems to us that a new norming technique will have to be devised for taxonomy-type tests which will have to take into account

the dependency noted above and to reconcile the fact that scores from taxonomy-type tests are partially ipsative and partially normative.

Reliability

The interpretation of the reliability data is also somewhat clouded because of the nature of the score distributions and the method of calculation which was used.

The apparently ideal method of determining reliability would have been test-retest but its use was precluded for two reasons. First, we did not have access to students for additional testing. Second, it is quite probable that there would be marked changes in the process levels of the items on the second administration due to learning during and memory of the first administration. Whether it would be justifiable to compute test-retest reliability on this basis is a question which can be answered only by the results of subsequent empirical study. The use of alternate forms; e.g., Atomic Structure and Glaciers, as a method of determining reliability would have been illogical in this study because one purpose of the study was to determine the generality of processes over specific contents. Because of these facts, the KR 20 was regarded as the most appropriate technique of those which were available.

Nevertheless, the meaningfulness of the coefficients which are presented for the first four levels are somewhat questionable because of the nature of the score distributions. Because difficulty should increase as level increases, it will probably be the case that the magnitude of internal consistency coefficients will decrease as level increases. Consequently, it seems that the proper evaluation of the obtained coefficients should be in terms of a set of level coefficients which would be theoretically generated on the bases of the increments in difficulty from one level to the next. Unfortunately, we were unable to develop this set of theoretical coefficients.

PRECEDING PAGE BLANK-NOT FILMED

VI. ANALYSIS AND PRESENTATION OF RESULTS

This section describes the procedures for and results of testing the hypotheses. The first part deals with the hypothesis of hierarchical structure; the second part deals with the hypothesis of generality of process over content; and the third part deals with the psychological structure of the taxonomy-type tests.

A. Hierarchical Structure

The *Taxonomy* is claimed to be hierarchical according to complexity from Knowledge through Evaluation, and to be cumulative in that any level of process includes the processes of all lower levels. As noted in Section II, this feature is considered to be an entry point for one who wishes to determine the construct validity of the *Taxonomy*. If the hierarchical structure cannot be demonstrated, then the taxonomy is of limited usefulness.

Procedure

To examine the cumulative hierarchical structure, it was hypothesized that data from the taxonomy-type tests should have the following characteristics: (a) the mean scores on subtests decrease as the taxonomic levels of the subtests increase; and (b) an intercorrelation matrix of taxonomic level scores exhibit simplex structure when judged in terms of the hypothesized structure.

Analysis of mean scores. To examine the hypothesis that mean scores decrease as level of complexity increases, the mean score for each level on each form in the battery was computed for each grade, nine through twelve. The approximate numbers of students contributing to each mean were 1,200 for grade nine, 1,400 for grade ten, 1,300 for grade eleven, and 1,200 for grade twelve.

Analysis of simplex structure. According to Guttman's simplex theory, if a group of tests differ only in complexity, then this fact will be observable in a matrix of intercorrelations derived from scores from the tests. According to the theory, as stated earlier in Section I, if a perfect simplex exists, then

$$r_{ik} = r_{ij}r_{jk} \text{ where } i < j < k.$$

If this condition is satisfied, the correlation matrix will be such that the largest values will lie along the principal diagonal, followed by the next larger values in the adjacent diagonals, with the smallest values in the upper right and lower left corners.

Kaiser (1962) proposed a method of scaling a simplex through a least squares analysis procedure. By applying his method to all possible orders of variables in a matrix, one can arrive at a best order. In addition, he proposes a method for calculating q^2, a statistic which indicates the goodness of fit of the data to the model. As the sum of squares accounted for by a given order approaches the total sum of squares, q^2 approaches unity. Therefore, the best order would be one which yields the maximum q^2 statistic.

Twenty 6 x 6 intercorrelation matrices were analyzed for simplex structure, one for each grade for each form and one for each form for all grades combined. Correlations were corrected for attenuation prior to the analysis. In each analysis, the Knowledge subtest was held as level 1 and the other five subtests were permuted. Therefore, 120 analyses were run for each form-grade combination. The decision to set arbitrarily the Knowledge subtest as level 1 was based on the belief that it was the least complex process because all relevant content to be used in responding to the Knowledge items was available to the student during the time he took the test.

Results

Mean scores. The results of the analysis of mean scores appear in Table 17. When considering them the reader should remember that the chance means on the Knowledge, Comprehension, Application, and Analyses subtests are five and that the chance means on Synthesis and Evaluation subtests are indeterminate because their items were in a free-response format.

The tabular data indicate that the predicted pattern of means is upheld at all grade levels on the Lisbon Earthquake and Stages of Economic Growth forms. The patterns of means for the two forms, Atomic Structure and Glaciers, which are based on science content, contain several departures from the predicted patterns. For each form and at every grade level, the Evaluation means are higher than the Synthesis means, and for the Glaciers form at all grade levels the Application means are higher than the Comprehension means.

TABLE 17

Comparison of Mean Performance on Florida Taxonomy Battery

Atomic Structure	Grade			
Process	9	10	11	12
Knowledge	13.0	14.0	14.0	15.3
Comprehension	6.9	8.2	8.5	8.9
Application	6.4	7.8	8.2	8.7
Analysis	5.2	6.3	6.8	7.0
Synthesis	.5	.9	1.2	1.4
Evaluation	1.2	1.6	1.7	2.1
Glaciers				
Process	9	10	11	12
Knowledge	15.7	16.4	17.1	17.4
Comprehension	7.7	8.9	9.7	9.9
Application	8.5	9.5	10.6	10.7
Analysis	6.3	7.2	7.6	7.9
Synthesis	3.1	3.8	4.0	4.0
Evaluation	4.2	4.6	5.2	5.6
Lisbon Earthquake				
Process	9	10	11	12
Knowledge	15.9	16.9	17.3	17.8
Comprehension	10.4	11.7	12.5	13.4
Application	8.5	9.9	10.6	11.4
Analysis	6.8	8.0	8.6	9.5
Synthesis	2.4	3.6	3.9	4.4
Evaluation	1.9	2.3	2.8	3.0
Stages of Economic Growth				
Process	9	10	11	12
Knowledge	14.3	15.1	15.9	16.2
Comprehension	8.2	9.4	10.2	10.7
Application	8.1	9.2	10.0	10.6
Analysis	6.2	7.1	7.7	8.1
Synthesis	1.2	1.8	2.4	2.4
Evaluation	.7	1.0	1.2	1.4

The extent of similarity between these patterns of means and the predicted pattern is regarded as generally supporting the hypothesis, but that support is not univocal because the science forms contained several systematic inconsistencies. But on balance, it appears that the data lend more support to accepting the hypothesis than rejecting it.

Simplex structure. Original correlations and correlations corrected for attenuation for each grade appear in Table 18, and comparable coefficients for all grades combined appear in Table 19. A summary of results of the simplex analyses for the Atomic Structure and Glaciers tests appears in Table 20 and the summary for the Lisbon Earthquake and Stages of Economic Growth tests appears in Table 21.

The simplex analyses reveal that for only the Lisbon Earthquake form does the taxonomic arrangement according to complexity correspond to the empirically determined order of complexity at each of the grade levels and over grades. The Atomic Structure form contains only one discrepancy between the theoretical and actual order of complexity. It occurs only at the ninth-grade level where the empirical ordering places Synthesis and Evaluation between Knowledge and Comprehension. The Stages of Economic Growth form contains only two discrepancies. In grade nine Synthesis is empirically placed between Knowledge and Comprehension; and in grade twelve, Synthesis and Evaluation are reversed. The Glaciers form reveals the same empirical pattern at all grade levels; Synthesis falls in that order between Knowledge and Comprehension.

The hypothesis of simplical structure appears not to be supported or rejected by the results. There is no statistical test of significance with which to reach a summary judgment about whether there is a non-chance relationship between the empirical order and the theoretical order. In the absence of it and for reasons that are presented subsequently, the evidence is regarded in the main as more confirming than disconfirming of the hypothesis.

Discussion of Results

The evidence about the hypotheses which was provided by the patterns of means and the simplical arrangements according to complexity is such that no definite assertion can be made about whether or not the taxonomic processes are arranged in hierarchical order. But the evidence and other

TABLE 18

Original Correlations and Correlations
Corrected for Attentuation*

Grades 9-12

	Grade 9						Grade 10						Grade 11						Grade 12					
Atomic Structure	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev	K	C	Ap	An	Sy	Ev
K.		371	322	120	324	363		528	483	364	401	374		581	566	440	421	380		464	491	361	406	361
C	498		556	485	164	268	694		660	604	401	362	763		775	693	428	368	610		731	683	435	305
Ap	422	302		443	214	279	634	952		571	440	346	744	999		695	454	336	644	999		662	457	315
An	165	732	668		064	121	500	912	862		367	300	605	999	999		378	322	496	999	999		393	268
Sy	382	215	273	086		375	465	510	560	489		449	488	545	577	504		389	471	553	582	524		396
Ev	436	353	367	167	436		449	478	456	414	522		456	486	442	445	453		433	403	415	371	461	
			N = 1184						N = 1119						N = 1258						N = 1196			
Glaciers																								
K		380	437	227	275	391		496	507	361	306	437		506	532	354	278	373		456	508	373	254	361
C	512		631	598	312	287	667		689	573	386	392	681		690	600	352	366	613		709	588	335	374
Ap	567	897		578	387	343	656	980		593	408	406	688	980		606	389	385	658	999		610	382	349
An	342	990	920		289	198	544	949	945		326	284	533	993	966		286	297	563	973	971		[illegible]	295
Sy	359	447	532	465		234	399	553	562	523		296	303	504	536	460		223	331	480	526	533		
Ev	500	402	463	310	318		559	550	548	447	406		477	514	520	468	304		462	525	472	464	385	
			N = 1223						N = 1131						N = 1303						N = 1186			
Lisbon Earthquake																								
K		544	517	418	481	318		565	548	485	482	341		567	550	494	444	318		526	544	497	435	300
C	725		642	573	475	304	753		721	671	557	393	756		717	678	538	384	700		686	661	474	301
Ap	715	898		565	498	315	759	999		684	538	356	761	999		704	563	390	752	958		652	489	286
An	581	804	824		441	273	673	941	996		550	381	686	951	999		542	381	690	927	950		525	327
Sy	656	654	733	633		392	656	767	789	789		439	605	741	805	778		487	593	653	699	754		380
Ev	406	392	422	366	517		435	507	476	512	578		403	495	522	512	642		383	388	383	439	501	
			N = 1194						N = 1451						N = 1298						N = 1192			
Stages of Economic Growth																								
K		496	482	248	368	319		532	508	348	370	361		483	500	307	386	337		463	464	341	330	330
C	633		698	578	426	413	679		733	648	490	455	617		772	705	517	471	591		739	697	466	459
Ap	621	946		522	444	421	655	993		580	443	420	644	999		626	514	463	598	999		651	468	433
An	349	855	780		300	272	490	959	865		389	378	432	999	934		423	404	480	999	971		421	405
Sy	457	556	584	437		318	458	638	582	558		397	478	674	676	607		412	410	608	615	604		385
Ev	414	565	580	408	422		469	621	579	569	527		438	643	638	607	547		429	627	597	609	510	
			N = 1199						N = 1451						N = 1283						N = 1194			

*Values above diagonals are original correlation coefficients; Those below the diagonal are corrected for attenuation. Decimal points have been omitted.

TABLE 19

Original Correlations and Correlations Corrected for Attentuation*

All Grades Together

Atomic Structure

	K	C	Ap	An	Sy	Ev
Knowledge		514	496	358	412	386
Comprehension	675		712	648	415	355
Application	651	999		632	450	348
Analysis	493	980	955		371	294
Synthesis	478	529	573	494		426
Evaluation	463	468	459	406	496	
N = 5057						

Glaciers

	K	C	Ap	An	Sy	Ev
Knowledge		481	513	351	298	409
Comprehension	648		698	608	372	382
Application	664	992		614	411	394
Analysis	529	999	978		331	295
Synthesis	389	533	567	532		277
Evaluation	627	535	532	463	376	
N = 5143						

Lisbon Earthquake

	K	C	Ap	An	Sy	Ev
Knowledge		577	564	498	486	340
Comprehension	760		719	675	551	379
Application	781	999		684	562	369
Analysis	692	947	997		558	376
Synthesis	663	759	804	801		453
Evaluation	434	488	494	505	597	
N = 5135						

Stages of Economic Growth

	K	C	Ap	An	Sy	Ev
Knowledge		512	503	337	385	353
Comprehension	654		753	680	509	472
Application	655	999		619	497	455
Analysis	474	999	924		424	397
Synthesis	477	664	654	610		409
Evaluation	459	645	628	597	543	
N = 5059						

*Values above diagonals are original correlation coefficients; those below the diagonal are corrected for attenuation. Decimal points have been omitted.

TABLE 20

Results of Simplex Analyses

Atomic Structure

	Grades				
	9	10	11	12	All grades
Maximum Obtained q^2 Value	.89	.92	.92	.92	.91
Order of Process Levels Leading to Maximum q^2 Value	1-5-6-2-3-4	1-2-3-4-5-6	1-2-3-4-5-6	1-2-3-4-5-6	1-2-3-4-5-6
Deviation of q^2 for Hypothesized Order from Maximum Obtained	.17	.00	.00	.00	.00

Glaciers

	Grades				
	9	10	11	12	All grades
Maximum Obtained q^2 Value	.90	.91	.89	.89	.95
Order of Process Levels Leading to Maximum q^2 Value	1-6-2-3-4-6	1-6-2-3-4-5	1-6-2-3-4-5	1-6-2-3-4-5	1-6-2-3-4-5
Deviation of q^2 for Hypothesized Order from Maximum Obtained	.08	.07	.08	.02	.13

TABLE 21

Results of Simplex Analyses

Lisbon Earthquake

	Grades 9	10	11	12	All grades
Maximum Obtained q^2 Value	.95	.986	.995	.983	.991
Order of Process Levels Leading to Maximum q^2 Value	1-2-3-4-5-6	1-2-3-4-5-6	1-2-3-4-5-6	1-2-3-4-5-6	1-2-3-4-5-6
Deviation of q^2 for Hypothesized Order from Maximum Obtained					

Stages of Economic Growth

	Grades 9	10	11	12	All grades
Maximum Obtained q^2 Value	.86	.90	.91	.92	.91
Order of Process Levels Leading to Maximum q^2 Value	1-5-2-3-4-6	1-2-3-4-5-6	1-2-3-4-5-6	1-2-3-4-5-6	1-2-3-4-5-6
Deviation of q^2 for Hypothesized Order from Maximum Obtained	.03	.00	.00	.01	.00

information about the tests lead to the belief that there is more confirmation than disconfirmation. The reasons for this are as follows.

First, only two kinds of discrepancies in means occurred; Application and Comprehension were reversed at all grade levels on one form, and on that form and the other science-content form Synthesis and Evaluation were reversed at all grade levels. With regard to the former, the average difference between the two means was about .7 and no pair of the means differed by as much as 1.0. The item analyses data, which appears in Appendix B, reveal that two of the Comprehension items were exceedingly difficult. If the reader attempts to solve these two items then he will probably have to read the content passages in the test booklet several times before he locates the appropriate information. The information is there but it is very brief and it is not conspicuously located. There is little doubt that had two other Comprehension items been chosen for inclusion in the passage then this reversal would not have occurred.

The reversal of Synthesis and Evaluation means on the two science-content forms at all grade levels requires a much more involved explanation. First, the reader will recall from earlier sections that we had least confidence in the technical quality of the Synthesis and Evaluation items, that our misgivings were greater about the Evaluation items than the Synthesis items because some of the former ones required applying external criteria which were not supplied with the relevant content, and that Synthesis and Evaluation items were cast in a free-response format after we utterly failed to write them in the multiple-choice format. Second, the means reveal that collectively the science forms were more difficult than the social science forms. Third, student reactions about the difficulty of the Atomic Structure test caused us, as the reader will recall, never to administer it as the first one of the several forms which were to be administered.

Consequently, the explanation of the reversal of Synthesis and Evaluation means on the two forms is as follows: The principles, concepts, and generalizations which appear in the science forms, especially in the Atomic Structure form, were much more difficult to grasp and use than were the same entities in the social science forms. This had the net effect of relatively decreasing performance on the Synthesis and Evaluation items, but the effect was somewhat cushioned on the Evaluation items because at least half of them required students to use external criteria in reaching an item response. We believe that these "external criteria" were largely "common sense," an attribute in which the students appeared not to be lacking.

This argument is partially supported by the results of the simplex analyses which show that according to empirically determined complexity both Synthesis and Evaluation are placed between Knowledge and Comprehension at grade nine on Atomic Structure; that Evaluation was placed between Knowledge and Comprehension at every grade level on the Glaciers form; and that all patterns on Lisbon Earthquake were as predicted. The argument is somewhat weakened because Synthesis was placed between Knowledge and Comprehension at grade nine and Synthesis and Evaluation were reversed at grade twelve on the Stages of Economic Growth form.

Although it is rather late to suggest that we have qualms about the rationale for the hypothesis of decreasing mean performance as process level increases, it seems appropriate to invoke Guttman's contention that item complexity and item difficulty are independent attributes. (Guttman, 1954, p. 283)

> ... It is entirely possible to make an addition test far more difficult than a subtraction test. For example, for a given population, select a set of ten addition items all of which were missed by at least 70% of the population. The mean total score on such a test will then be less than 3. Select ten subtraction items all of which are answered correctly by at least 70% of the population. The mean total score on subtraction will then be at least 7. Hence the population mean on subtraction is greater than for addition, or this particular subtraction test is less difficult than this particular addition test. However, the intercorrelation between these two tests can be precisely the same as for tests in which subtraction is made more difficult than addition. Furthermore, the intercorrelations with other tests in a simplex can be essentially the same, no matter if orders of difficulty are reversed (p. 285).

In this same vein, Crawford (1966) recently reported empirical findings, based on taxonomic data, that complexity and difficulty were unrelated over six hierarchical processes as described in an adaptation of the Taxonomy.

Our rationale for the relevant hypothesis was based on the particular notion that if the relevant test content were hitherto unfamiliar to students and that if all students were supplied with it prior to the test and had equal access to it during the test, then knowledge of content would be a

constant factor over all levels of process with the consequence that item difficulty would increase as complexity of items increased. With just some exceptions, the pattern of means of support this rationale; however, the simplex data assail it because of the great frequency with which the most difficult tests, Synthesis and Evaluation, were placed according to complexity between Knowledge and Comprehension. The force of this contradictory evidence is somewhat lessened due to the different item formats and hypothetical chance scores, but the point still stands.

The results of the simplex analyses are somewhat contradictory, but they are regarded with some optimism. On one form they corroborated the imputed order of the taxonomy, and on the other three forms the predicted pattern occurred except for placement of Synthesis and Evaluation both of which consisted of questionable items. Although the hypothesis remains without verification, the results suggest that further work on item construction is needed before the hypothesis can validly be subjected to testing. However, that aspect of the Evaluation process which requires the use of external criteria might consistently lead to the misplacement of it unless some method is devised by which to make mastery of the external criteria a common possession of all testees.

B. Transcendence of Process

It is claimed that the processes described in the *Taxonomy* are general over content. The response to a test item is conceptualized as the interaction of content and process. When a student identifies the appropriate content, then he must use the relevant process to manipulate the content to produce a response to the item. Accordingly, Application in mathematics, Application in science, and Application in geography are identical with respect to process and differ only in the contents which are manipulated.

Procedure

To examine the generality of process, it was hypothesized that data from taxonomy-type tests would have the following characteristics: (a) a factor analysis of the level scores from the four taxonomy-type tests will have the structure shown in the following scheme:

	Factors					
Processes	I	II	III	IV	V	VI
Knowledge	X					
Comprehension	X	X				
Analysis	X	X	X			
Application	X	X	X	X		
Synthesis	X	X	X	X	X	
Evaluation	X	X	X	X	X	X

An "X" indicates the expected presence of important factor loadings for a level score for each of the four tests. The hierarchical structure of the processes is shown by the fact that after Knowledge each level score contains all the factors contained by the preceding level score plus one additional factor; and (b) the intercorrelations of scores from the same subtest for different contents will have circumplex structure.

<u>Factor analysis.</u> The scores from the taxonomy-type tests were factor analyzed. Each of the four forms contained six subtests, thereby yielding twenty-four process scores. A 24 x 24 matrix of Pearson product-moment coefficients, with unities in the principal diagonal, was analyzed by the principal axis method for each grade, nine through twelve. Twenty-four factors were extracted from each matrix. Although Kaiser's rule indicated that only four factors were present in each matrix, six were retained for interpretation because that number had been hypothesized.

<u>Circumplex analysis.</u> According to Guttman's circumplex theory, if a group of tests differ only in content, an intercorrelation matrix derived from the scores will have two unique characteristics; the column totals will be equal and each row of the table will have the same entries as the preceding row, but moved one space to the right, and the end one moving to the beginning.

Thirty correlation matrices were constructed, one for each process level for each grade and one for each level for all grades combined. Correlations were corrected for attenuation prior to the analysis.

Results

Factor analysis. The usual methods of factor analysis prescribe a rotational procedure which seeks a simple structure and, in the process, distributes the variance of a general factor over other factors. Because the notion of simple structure is negated by the hypothesized factor structure for the cumulative hierarchy, the decision was made to attempt to interpret the unrotated factor matrices. Although the interpretation of unrotated matrices is not very desirable, it seems to be one of the few avenues of interpretation which was available. The major disadvantage in dealing with unrotated factors is that all factors are bipolar except the first one.

The unrotated factor matrices appear in Tables 22 through 25.

To facilitate interpretation, a two-way table was constructed for factors II through VI (factor I is a general factor) for each grade level. The rows in the tables represent the four test forms, and the six columns represent the processes. The cell entries are factor loadings. These appear as Tables 26 through 29.

For each of the four grade levels, a general factor emerged. It accounted for 39%, 49%, 46% and 43% of the variance for grades nine through twelve, respectively. For the ninth grade, Application for the form Stages of Economic Growth has no loading on the general factor. All other tests at all grade levels had high--or in a few cases, moderate--loadings on the general factor.

When interpreting the factors, the pattern of signs and the magnitude of the loadings were taken into account. For a given factor, if all four measures for a certain process had non-zero loadings (approximately greater than .10 since the factors were unrotated) and were of the same sign, then that process was considered to be represented on that factor. In some instances the process was identified as a factor when only three measures of it were like signed and had non-zero loadings; however, this was never done when the fourth measure had a moderate loading with a sign unlike the other three. In identifying a factor as being primarily related to a specific content the same criteria were adopted except that four of the six measures were required to have like signs and non-zero content.

In the matrices for all four grade levels some of the factors can be identified as process factors; i.e., the factor

TABLE 12

Unrotated Factor Matrix

Taxonomy-Type Tests

Grade 9

Variable	FACTORS I	II	III	IV	V	VI
Atomic Structure						
K	-685	240	071	-372	091	-015
C	-567	-373	-438	-222	027	-081
Ap	-557	-338	-408	-187	-002	-112
An	-343	-546	-487	-100	036	-011
S	-484	251	-029	259	-305	-349
E	-546	296	-289	175	-073	-206
Glaciers						
K	-681	293	050	-337	246	-006
C	-697	-302	245	066	070	-011
Ap	-743	-251	224	125	101	-012
An	-520	-466	337	095	029	-168
S	-437	-117	240	187	300	-538
E	-535	180	-231	169	343	-060
Lisbon Earthquake						
K	-665	221	121	-385	-077	-107
C	-745	-067	174	-181	-201	037
Ap	-744	-059	175	-154	-280	-037
An	-663	-032	124	-013	-375	149
S	-673	169	-108	068	-271	-133
E	-500	336	-354	139	-125	049
Economic Growth						
K	-632	405	056	-224	275	148
C	-776	-059	064	152	089	281
Ap	-076	-047	033	121	154	227
An	-608	-238	042	205	-075	375
S	-553	217	-018	300	-050	086
E	-527	135	-111	386	172	110

TABLE 23

Unrotated Factor Matrix

Taxonomy-Type Tests

Grade 10

Variable	FACTORS I	II	III	IV	V	VI
Atomic Structure						
K	697	180	281	-312	-043	-096
C	706	-085	-204	-416	-141	130
Ap	699	-119	-254	-372	-122	-025
An	609	-122	-363	-396	-174	198
S	586	302	-320	027	-092	-218
E	550	438	-222	-002	097	-088
Glaciers						
K	689	188	460	-163	-023	-168
C	766	-250	013	011	-040	-145
Ap	757	-268	037	026	-019	-211
An	593	-421	-002	049	-007	-396
S	527	-018	-228	292	-170	-396
E	612	386	-105	000	166	-082
Lisbon Earthquake						
K	650	152	455	089	-244	004
C	762	-128	141	196	-263	225
Ap	776	-138	165	195	-162	157
An	741	-082	047	201	-188	276
Sy	721	189	-066	206	-118	069
E	529	353	-156	187	-160	213
Economic Growth						
K	633	231	411	-232	269	-006
C	790	-204	010	028	302	070
Ap	753	-180	078	034	311	056
An	651	-349	-092	110	317	163
S	633	160	-224	169	202	-144
E	583	143	-128	104	327	237

TABLE 24

Unrotated Factor Matrix

Taxonomy-Type Tests

Grade 11

Variable	I	II	III	IV	V	VI
	FACTORS					
Atomic Structure						
K	-700	186	220	241	183	-246
C	-757	017	-281	327	041	-230
Ap	-761	-043	-299	288	067	-238
An	-662	-030	-371	313	024	-251
S	-598	301	-177	-320	-022	-254
E	-489	567	-107	019	079	156
Glaciers						
K	-630	036	485	122	270	047
C	-750	-257	-023	-012	268	118
Ap	-756	-279	036	-087	279	130
An	-608	-372	-084	-080	361	224
S	-440	-060	027	-534	355	-387
E	-565	427	-019	-023	228	331
Lisbon Earthquake						
K	-639	029	482	100	-143	-094
C	-774	-195	153	-109	-194	004
Ap	-772	-173	153	-173	-167	-029
An	-739	-150	086	-175	-254	-070
S	-719	165	-040	-290	-188	-135
E	-600	405	-174	039	-013	167
Economic Growth						
K	-612	220	449	290	-113	-046
C	-819	-204	-086	115	-152	136
Ap	-807	-164	020	080	-124	146
An	-679	-288	-237	053	-259	190
S	-638	147	-032	-247	-253	050
E	-607	195	-128	-095	-014	286

TABLE 25

Unrotated Factor Matrix

Taxonomy-Type Tests

Grade 12

Variable	FACTORS I	II	III	IV	V	VI
Atomic Structure						
K	657	-233	-022	-092	-393	033
C	694	274	365	-158	-203	-032
Ap	734	283	313	-143	-203	-031
An	641	324	418	-166	-107	-050
S	602	-154	338	133	046	139
E	500	-366	264	248	-017	-382
Glaciers						
K	640	-205	-327	-036	-338	200
C	755	208	-078	106	-171	-012
Ap	759	237	-101	172	-165	-000
An	609	296	-075	276	-193	043
S	457	128	123	400	132	564
E	553	-429	111	188	005	-028
Lisbon Earthquake						
K	643	-151	-367	035	-134	006
C	743	145	-268	147	099	-174
Ap	716	109	-315	202	102	-177
An	738	072	-237	167	213	-139
S	679	-273	036	164	249	163
E	517	-407	240	135	-007	-169
Economic Growth						
K	514	-453	-152	-445	-181	207
C	771	079	-116	-329	291	-039
Ap	751	045	-124	-332	182	036
An	702	160	-091	-028	279	-137
S	602	-094	155	-120	340	345
E	610	-193	198	-064	193	-175

TABLE 26

Factor Loadings for

Factors II through VI

Grade 9

Test Form	K	C	Ap	An	S	E	Factor
			Process Levels				
As	+24	-37	-34	-55	+25	+30	II
G	+29	-30	-25	-47	-12	+18	
L	+22	-07	-06	-03	+17	+34	
EG	+40	-06	-05	-24	+22	+14	
As	+07	-44	-41	-49	-03	-29	III
G	+05	+25	+22	+34	+24	-23	
L	+12	+17	+18	+12	-11	-35	
EG	+06	+06	+03	+04	-02	-11	
As	-37	-22	-19	-10	+26	+17	IV
G	-34	+06	+13	+10	+19	+17	
L	-39	-18	-15	-01	+07	+13	
EG	-22	+15	+12	+21	+30	+38	
As	+09	+03	-00	+04	-31	-07	V
G	+25	+07	+10	+03	+30	+34	
L	-08	-20	-29	-38	-27	-13	
EG	+28	+09	+15	-08	-05	+17	
As	-02	-08	-11	-01	-35	-21	VI
G	-01	-01	-01	-17	-54	-06	
L	-11	+04	-04	+15	-13	+05	
EG	+15	+28	+23	+38	+09	+11	

TABLE 27

Factor Loadings for

Factors II through VI

Grade 10

Test Form	K	C	Ap	An	S	E	Factor
			Process Levels				
As	+18	-09	-12	-12	+30	+44	II
G	+18	-25	-27	-42	-02	+39	
L	+15	-13	-14	-08	+19	+35	
EG	+23	-20	-18	-35	+16	+14	
As	+28	-20	-25	-36	-32	-22	III
G	+46	+01	+03	-00	-23	-11	
L	+46	+14	+17	+05	-07	-16	
EG	+41	+01	+08	-09	-22	-13	
As	-31	-42	-37	-40	+03	-00	IV
G	-16	+01	+03	+05	+29	+00	
L	+09	+20	+20	+20	+21	+19	
EG	-23	+03	+03	+11	+17	+10	
As	-04	-14	-12	-17	-09	+10	V
G	-02	-04	-02	-01	-17	+17	
L	-24	-26	-16	-19	-12	-16	
EG	+27	+30	+31	+32	+20	+33	
As	-10	+13	-03	+20	-22	-09	VI
G	-17	-15	-21	-40	-40	-08	
L	+00	+23	+16	+28	+07	+21	
EG	-01	+07	+06	+16	-11	+24	

TABLE 28

Factor Loadings for

Factors II through VI

Grade 11

Test Forms	K	C	Process Level Ap	An	S	E	Factor
As	+19	+02	-04	-03	+30	+57	II
G	+04	-26	-28	-37	-06	+43	
L	+03	-20	-17	-15	+17	+41	
EG	+22	-20	-16	-29	+15	+20	
As	+22	-28	-30	-37	-18	-11	III
G	+49	-02	+04	-08	+03	-02	
L	+48	+15	+15	+09	-04	-17	
EG	+45	-09	+02	-24	-03	-13	
As	+24	+33	+29	+31	-32	+02	IV
G	+12	-01	-09	-08	-53	-02	
L	+10	-11	-17	-18	-29	+04	
EG	+29	+12	+08	+05	-25	-10	
As	+18	+04	+07	+02	-02	+08	V
G	+27	+27	+28	+36	+36	+23	
L	-14	-19	-17	-25	-19	-01	
EG	-01	-15	-12	-26	-25	-01	
As	-25	-23	-24	-25	-25	+16	VI
G	+05	+12	+13	+22	-39	+33	
L	-09	+00	-03	-07	-14	+17	
EG	-05	+14	+15	+19	+05	+29	

TABLE 29

Factor Loadings for

Factors II through VI

Grade 12

Test Form	K	C	Ap	An	S	E	Factor
			Process Levels				
As	-23	+27	+28	+32	-15	-37	II
G	-21	+20	+24	+30	+13	-43	
L	-15	+15	+11	+07	-17	-41	
EG	-45	+08	+05	+16	-09	-19	
As	-02	+36	+31	+42	+34	+26	III
G	-33	-08	-10	-08	+12	+11	
L	-37	-27	-32	-24	+03	+24	
EG	-15	-12	-12	-09	+16	+20	
As	-09	-16	-14	-17	+13	+25	IV
G	-04	+11	+17	+28	+40	+19	
L	+04	+15	+20	+17	+16	+14	
EG	-45	-33	-33	-03	-12	-06	
As	+39	+20	+20	-11	+04	-02	V
G	+34	-17	-17	-19	+13	+01	
L	-13	+10	+10	+21	+25	-01	
EG	-18	+23	+18	+28	+34	+19	
As	+03	-03	-03	-05	+14	+38	VI
G	+20	-01	-00	+04	+56	-03	
L	+01	-17	-18	-14	+16	-17	
EG	+21	-04	+04	-14	+35	-18	

loadings pervade across the contents. At all grade levels at least one factor of the six extracted is a content factor according to the criteria described above and in some instances mixed factors containing variance due to process and content were obtained.

At grade nine two factors, in addition to the general factor, were found to be primarily process factors. Two were mixed factors and one primarily involved content. Factor II is Knowledge and Evaluation vs Analysis, and Factor IV is Knowledge vs Synthesis and Evaluation. The mixed factors are as follows: III is Evaluation plus Atomic Structure and VI, Synthesis vs Economic Growth. Factor V is a Glaciers factor.

At grade ten two process, one content, and two mixed factors emerged. Factor II is Knowledge, Synthesis, and Evaluation vs Comprehension, Application and Analysis. Factor III is Knowledge vs Synthesis and Evaluation. The mixed factors are: IV is Knowledge vs Synthesis and Lisbon vs Atomic Structure; and VI is Synthesis and Glaciers. The content factor is number V which is defined by Economic Growth vs Lisbon.

At grade eleven three process factors, one content factor, and one mixed factor appeared. The process factors are as follows: Factor II is Comprehension, Application and Analysis vs Synthesis and Evaluation; Factor III is Knowledge vs Evaluation; and Factor IV is Knowledge vs Synthesis. The content factor is V, Glaciers vs Lisbon and Economic Growth. The mixed factor is VI, Synthesis vs Economic Growth.

At grade twelve two process and three mixed factors emerged in addition to the general factor. Factor II is Knowledge and Evaluation vs Comprehension, Application, and Analysis. Factor V is a relatively pure Synthesis factor. The mixed factors are as follows: Factor III is Knowledge vs Synthesis and Evaluation plus specific content on Atomic Structure; Factor IV is Evaluation plus Lisbon vs Economic Growth, and Factor VI is Synthesis vs Lisbon.

Circumplex analysis. Thirty correlation matrices were constructed, one for each process level for each grade and one for each process for all grades combined. These matrices appear in Tables 30 through 34. While none of these matrices can be classified as a perfect circumplex in the sense that all column totals are exactly equal, many of them approach this state. As an aid in interpreting the goodness of fit of a matrix of similar process correlations to a perfect circumplex, the smallest column total was subtracted from

the largest column total for each matrix and a distribution of these differences was made. This distribution was found to be highly positively skewed. It was estimated that a symmetrical distribution about the median would include a range of differences from about .01 to about .39, and that differences of .40 and beyond probably indicate that the associated matrix is not a circumplex. For purposes of the present analysis then, matrices whose high-low column sum differences fall within the range of .01 to .39 are considered to be quasi-circumplexes.

For grade nine, only the matrices for Knowledge and Evaluation can be considered as quasi-circumplexes. For grade ten all matrices satisfy the criterion for the quasi-circumplex. For grades eleven and twelve only the matrices for Synthesis fail to do so. All process matrices for all grades combined satisfy the quasi-circumplex criterion with the exception of Synthesis. These interpretations of the circumplex analyses tend to support the theory underlying the taxonomy-type tests. However, it must be remembered that the cut off of .40 was established in a very arbitrary manner. Until either theoretical or empirical research concerning the true distribution of the high-low column sum differences is undertaken, these interpretations must be considered as highly tentative.

Discussion

The hypothesis tested by the factor analyses of the taxonomy-type tests and the circumplex analyses was that of transcendence of process. More specifically a hierarchical structure of the processes was postulated and it was predicted that the intercorrelations of scores from the subtests across forms would form a circumplex or quasi-circumplex.

The factor analyses offer some support for the hypothesis. Of the non-general factors, several at each grade level give an indication of being more highly determined by the processes than by the contents; moreover, a fairly high degree of consistency exists over grade levels with regard to the nature of the process factors. The following informal equations show the patterns of process loadings on factors over the grade levels.

$$F_{9,II} = K - Ap + E$$

$$F_{10,II} = K - C - An - Ap + S + E \text{ (reflected)}$$

$$F_{11,II} = -C - An - Ap + S + E$$

$$F_{12,II} = K - C - An - Ap + E \text{ (reflected)}$$

TABLE 30

Correlations Within Levels Across Forms

Grade 9

	K_1	K_2	K_3	K_4
K_1		719	702	723
K_2	719		750	759
K_3	702	750		591
K_4	723	759	591	
	2.144	2.228	2.043	2.073

	C_1	C_2	C_3	C_4
C_1		587	576	541
C_2	587		719	821
C_3	576	719		723
C_4	541	821	723	
	1.704	2.127	2.017	2.085

	Ap_1	Ap_2	Ap_3	Ap_4
Ap_1		567	544	552
Ap_2	567		803	784
Ap_3	544	803		723
Ap_4	552	789	723	
	1.663	2.159	2.070	2.064

	An_1	An_2	An_3	An_4
An_1		470	292	395
An_2	470		532	572
An_3	292	532		650
An_4	395	572	650	
	1.157	1.574	1.474	1.617

	S_1	S_2	S_3	S_4
S_1		280	472	400
S_2	280		358	309
S_3	472	358		553
S_4	400	309	553	
	1.152	.947	1.383	1.262

	E_1	E_2	E_3	E_4
Ev_1		482	487	445
Ev_2	482		452	468
Ev_3	487	452		434
Ev_4	445	468	434	
	1.414	1.402	1.373	1.347

TABLE 31

Correlations Within Levels Across Forms

Grade 10

	K_1	K_2	K_3	K_4
K_1		739	652	725
K_2	739		814	768
K_3	652	814		610
K_4	725	768	610	
	2.116	2.321	2.076	2.103

	C_1	C_2	C_3	C_4
C_1		791	723	751
C_2	791		829	858
C_3	723	829		754
C_4	751	858	754	
	2.265	2.478	2.306	2.363

	Ap_1	Ap_2	Ap_3	Ap_4
Ap_1		741	670	668
Ap_2	741		843	763
Ap_3	670	843		816
Ap_4	668	763	816	
	2.079	2.347	2.329	2.247

	An_1	An_2	An_3	An_4
An_1		548	616	624
An_2	548		650	732
An_3	616	650		727
An_4	624	732	727	
	1.788	1.930	1.993	2.083

	S_1	S_2	S_3	S_4
S_1		468	632	530
S_2	468		585	492
S_3	632	585		692
S_4	530	492	692	
	1.630	1.545	1.909	1.714

	E_1	E_2	E_3	E_4
E_1		634	463	455
E_2	634		527	565
E_3	463	527		432
E_4	455	565	432	
	1.552	1.726	1.422	1.452

TABLE 32

Correlations Within Levels Across Forms

Grade 11

	K_1	K_2	K_3	K_4
K_1		679	628	684
K_2	679		657	654
K_3	628	657		684
K_4	684	654	684	
	1.991	1.990	1.969	2.022

	C_1	C_2	C_3	C_4
C_1		829	728	575
C_2	829		806	872
C_3	728	806		833
C_4	575	872	833	
	2.132	2.507	2.367	2.280

	Ap_1	Ap_2	Ap_3	Ap_4
Ap_1		757	726	810
Ap_2	757		834	841
Ap_3	726	834		852
Ap_4	810	841	852	
	2.293	2.432	2.412	2.503

	An_1	An_2	An_3	An_4
An_1		653	694	751
An_2	653		713	737
An_3	694	713		756
An_4	751	737	756	
	2.098	2.103	2.163	2.244

	S_1	S_2	S_3	S_4
S_1		462	712	544
S_2	462		527	379
S_3	712	527		773
S_4	544	379	773	
	1.718	1.368	2.012	1.696

	E_1	E_2	E_3	E_4
E_1		610	535	459
E_2	610		617	561
E_3	535	617		579
E_4	459	561	579	
	1.604	1.788	1.731	1.599

TABLE 33

Correlations Within Levels Across Forms

Grade 12

K	K_1	K_2	K_3	K_4
K_1		615	594	574
K_2	615		706	598
K_3	594	706		473
K_4	574	598	473	
	1.783	1.919	1.773	1.645

	C_1	C_2	C_3	C_4
C_1		753	668	719
C_2	753		795	767
C_3	668	795		749
C_4	719	767	749	
	2.140	2.315	2.212	2.235

	Ap1	Ap_2	Ap_3	Ap_4
Ap_1		793	642	708
Ap_2	793		772	702
Ap_3	642	772		698
Ap_4	708	702	698	
	2.143	2.263	2.112	2.108

	An_1	An_2	An_3	An_4
An_1		661	590	737
An_2	661		723	681
An_3	590	723		804
An_4	737	681	804	
	1.988	2.065	2.117	2.222

	S_1	S_2	S_3	S_4
S_1		385	623	500
S_2	385		494	415
S_3	623	494		660
S_4	500	415	660	
	1.508	1.294	1.777	1.575

	E_1	E_2	E_3	E_4
E_1		525	482	488
E_2	525		576	531
E_3	482	576		486
E_4	488	531	486	
	1.495	1.632	1.544	1.505

TABLE 34

Correlations Within Levels Across Forms

All Grades

	K_1	K_2	K_3	K_4
K_1		705	667	689
K_2	705		767	726
K_3	667	767		629
K_4	689	726	629	
	2.051	2.198	2.063	2.044

	C_1	C_2	C_3	C_4
C_1		784	716	778
C_2	784		826	852
C_3	716	826		804
C_4	778	852	804	
	2.278	2.462	2.336	2.434

	Ap_1	Ap_2	Ap_3	Ap_4
Ap_1		754	843	737
Ap_2	754		848	803
Ap_3	843	848		822
Ap_4	737	803	822	
	2.334	2.405	2.413	2.362

	An_1	An_2	An_3	An_4
An_1		624	627	702
An_2	624		708	728
An_3	627	708		785
An_4	702	728	785	
	1.953	2.060	2.120	2.215

	S_1	S_2	S_3	S_4
S_1		451	672	547
S_2	451		548	443
E_3	672	548		773
E_4	547	443	773	
	1.670	1.442	1.993	1.763

	E_1	E_2	E_3	E_4
E_1		592	512	488
E_2	592		577	559
E_3	512	577		520
E_4	488	559	520	
	1.592	1.728	1.609	1.567

In each of these factors two or more of the subtests, Knowledge, Synthesis, and Evaluation are always in opposition to Comprehension, Application or Analysis.

A second factor that emerged for the four grades is stated informally below.

$$F_{9,IV} = K - S - E \text{ (reflected)}$$

$$F_{10,III} = K - S - E$$

$$F_{11,III} = K - E$$

$$F_{12,III} = K - S - E \text{ (reflected)}$$

At all grade levels, factors involving only Synthesis were present and for grades nine, ten, and twelve, a factor involving only Evaluation was found.

Almost all of these factors were mixed in the sense that the content of specific test forms was always involved in the factor in addition to the process. The presence of these content-process factors tends to discredit the hypothesis. Their appearance suggests highly complex content-process interactions which might also be implied by the conceptualization of an item response as an interaction of content and a cognitive process. It remains for further analyses to clarify the nature of these mixed factors.

The transcendence hypothesis is supported to some extent by the circumplex analyses but convincing evidence cannot be provided by this technique because the probability of obtaining perfect circumplexes from empirical matrices is extremely low. Research on the sampling distribution of the high-low column sum differences is needed before this technique will be of practical value.

The decision to extract principal axis factors of the processes (subtests) precluded determining whether the processes approximate a hierarchical structure which would constitute another test of the hypothesis which is dealt with in Part A of this section. The appearance of a strong general factor at each grade gives some credence to the hypothesis but the bipolar form of the subsequent factors precludes a comprehensive testing of it and the simplical structure. Because of this, testing the hierarchical hypothesis was done by inspecting patterns of means and the simplical structure. Factor analysis results might have been more pertinent but it is difficult to determine what techniques are most appropriate. Three possible approaches are mentioned below for consideration in future analyses.

(a) A bi-factor solution (Harman, 1960) enables grouping variables according to a hypothesized design. The factor analysis allows the design to be tested by giving evidence of the adequacy of the hypothesized grouping. This technique was not used because computer programs were not available and the analysis of four 24 x 24 matrices would be too time consuming to accomplish by hand.

(b) The square root method of factor analysis would be particularly appropriate because the resulting factor matrix is triangular in form as is the hypothesized hierarchical structure. This method, however, is especially sensitive to the communality estimates one uses and will break down if the Gramian qualities of the matrix are not preserved by the communalities. Nevertheless this method should probably be attempted in future analyses.

(c) The presence of content factors in the taxonomy-type tests might obscure the results of either of these suggested methods. A third method would be to retain the principal factor solution and to attempt to rotate it to a hypothesized solution which would take into account both the content and process factors. An approach of this kind might employ the "pattern quartimax" technique which was developed by Saunders (1960) or by the "Procrustes" method proposed by Hurley and Cattell (1962).

C. Psychological Structure

One goal which guided the taxonomists was to produce a classification scheme which would be psychologically meaningful. This section presents the procedures and results of studies designed to reveal the psychological structure of the cognitive processes.

Procedure

Several major objectives were set for these studies: (a) to determine the cognitive aptitudes, as defined by KIT tests, and their regression weights which are related to taxonomy level scores; (b) to determine the changes in regression equations for each level, for all contents, for each grade; (c) to determine the similarity of equations for the same level for each content, but calculated on different grade-groups; and (d) to examine the equations for each taxonomy level, considered over all contents and grade-groups.

Attaining these objectives should reveal the structure of the taxonomy as defined by more elemental cognitive aptitudes, whether aptitudes transcend content, whether the

psychological structure of the taxonomy is a function of grade-level, and whether systematic changes occur from level to level in the regression equations.

The following sections deal with the procedures used to develop cognitive factor scores for use in the regression analysis.

Obtaining factor scores. Thirty-seven cognitive factor tests from KIT were administered to students from grades nine through twelve. A listing of these tests and the factors they purport to measure appears in Table 35. The aptitude tests were chosen to provide measures of sixteen cognitive factors which were believed to be related to the processes measured by the taxonomy-type tests. The scores from one aptitude test for grade nine students were judged to be invalid so they were discarded.

For twenty-four of the tests, scores were available for Part I and Part II. In each such case, Part I - Part II correlations were computed to yield an alternate forms reliability estimate. Because total scores were to be used in the analyses, these coefficients were used to estimate total test reliability by means of the Spearman-Brown prophecy formula. The obtained reliability coefficients appear in Table 36.

For each grade separately, the reference variables were intercorrelated by means of a missing point program to maximize the number of students contributing to each coefficient. The minimum number contributing to any coefficient was more than 200 and the majority of coefficients were based on more than 250 students.

Each correlation matrix was then factor analyzed by the principle axis method, using squared multiple correlation coefficients as communality estimates. Rotational solutions were imposed on the factor matrix by the varimax method.

It was presumed that the rotation would yield sixteen common factors because the KIT battery contained tests purported to measure that many psychologically pure factors. However, the rotation yielded a large number of specific factors so the sixteen-factor solution was abandoned. A ten-factor rotation solution for the twelfth-grade data was satisfactory because it contained no specific factors. On this basis, a ten-factor solution was regarded as appropriate for describing the aptitude data from all other grade levels.

TABLE 35

Kit of Reference Tests Administered
and Factors Identification

	Test Name	Factors
Cf-1	Hidden Figures Test	Flexibility of Closure
Cf-2	Hidden Patterns Test	" " "
Cf-3	Copying Test	" " "
Cs-1	Gestalt Completion Test	Speed of Closure
Cs-2	Concealed Words Test	" " "
Fa-1	Controlled Associations Test	Associational Fluency
Fa-3	Associations IV	" "
Fe-2	Simile Interpretations	Expressional Fluency
Fe-3	Word Arrangements	" "
Fi-1	Topics Test	Ideational Fluency
Fi-3	Thing Categories Test	" "
I-1	Letter Sets Test	Induction
I-2	Locations Test	"
I-3	Figure Classification	"
Le-1	Estimation of Length Test	Length of Estimation
Le-2	Shortest Road Test	" " "
Le-3	Nearer Point Test	" " "
Ma-1	Picture-Number Test	Associative Memory
Ma-3	First and Last Names Test	" "
N-1	Addition Test	Number Facility
N-2	Division Test	" "
N-3	Subtraction and Multiplication Test	" "
O-1	*Plot Titles	Originality
O-2	Symbol Production	"
R-3	Ship Destination Test	General Reasoning
R-4	Necessary Arithmetic Operations	" "
Re-1	Gestalt Transformation	Semantic Redefinition
Re-2	Object Synthesis	" "
Rs-1	Nonsense Syllogisms Test	Syllogistic Reasoning
Rs-2	Logical Reasoning Test	" "
Rs-3	Inference Test	" "
Sep-1	Apparatus Test	Sensitivity to Problems
Sep-2	Seeing Problems	" " "
V-1	Vocabulary	Verbal Comprehension
V-3	Wide Range Vocabulary Test	" "
Xa-1	Match Problems II	Figural Adaptive Flexibility
Xa-3	Planning Air Manuevers	" " "

* Scores on this test were judged invalid for grade nine.

TABLE 36

Reliability Estimates for Twenty-four
KIT of Reference Tests

Test Code*	Sample Size	Reliability Estimate (Part I-Part II)	Reliability Estimate (Total Test)
Cf-1	1107	.780	.876
Cf-3	1098	.742	.852
Cs-1	1128	.750	.867
Cs-2	1110	.667	.800
Fe-2	1102	.202	.337
Fe-3	1047	.489	.657
Fi-1	1098	.745	.854
Fi-3	1052	.401	.573
I-2	1042	.476	.645
I-3	1094	.128	.226
Le-2	1104	.439	.610
Le-3	1124	.569	.725
Ma-1	1048	.534	.696
Ma-3	1095	.719	.836
N-1	1098	.808	.894
N-2	1100	.795	.886
N-3	1096	.831	.908
O-2	1076	.655	.791
R[illegible]-2	1090	.616	.763
Rs-1	1062	.211	.348
Sep-2	1074	.497	.664
Xa-1	1073	.651	.788
Xa-3	1073	.328	.494
Xs-3	1071	.224	.366

*Test code symbols are those associated with test names in Table 35.

Factor scores were computed through a linear regression technique in which the common-factor score for subject i was estimated by

$$f_{ij} = B_{j1}Z_{j2}Z_{i2} + \ldots + B_{jn}Z_{in} = M$$

where $j = 1, 2, 3, \ldots n$

and Z_{ik} for $k = 1, 2, 3 \ldots n$ are the obtained scores on the n tests, expressed in standardized score form. The B_{ij} values are beta weights and are determined by minimizing the sum of squares of the discrepancies between the hypothetical true factor, f'_{ij}, and the estimate, f_{ij}. This solution is a special case of the least square multiple regression equation (Harman, 1960). This technique yielded a set of ten KIT factor scores for each subject.

Regression analyses. Distributions of factor scores and taxonomy test scores were transformed so that each would have a mean of 50 and a standard deviation of 10. Because the regression analyses required complete data for each subject, a score of 50 was assigned to each missing data point. Each set of factor scores was used to predict the performance of students on each of the six subtests of each of the taxonomy forms. Thus twenty-four regression analyses were computed for each grade level. The regression equations were developed by standard methods.

Results

Kit of Reference Tests. A preliminary factor analysis of the twelfth-grade data indicated the presence of ten common factors. Because the analyses for all grades were to be as comparable as possible, ten factors were extracted for all four grades. The unrotated factor matrices appear in Tables 37 through 40; the unrotated varimax solutions appear in Tables 41 through 44. Tables 45 through 51a contain eight factors which appear to be the same over grade ten through twelve, six of which also appear at grade nine. Tables 52 through 55 contain the loadings for the remaining factors in grades 9 through 12.

Factor loadings of .30 or greater were arbitrarily considered to indicate an interpretable correlation between a factor and a test. The eight common factors were named Number Facility, Associative Memory, Speed of Closure, Fluency, Verbal Reasoning, Flexibility of Closure, Reasoning, and Fluency in Making Verbal Transformations. The composition

of each of these factors is not precisely the same at all grade levels but it is sufficiently similar from grade to grade to warrant regarding each as common to all grade levels. We do not believe that ignoring the slight differences in composition which do exist will result in invalid interpretations of the general tendencies.

Number Facility (Table 45), Associative Memory (Table 46), and Speed of Closure (Table 47) factors are present in all four grade analyses and are defined by the tests indicated for them in the KIT Manual.

The Fluency factor (Table 48) is present at all grade levels and it accounts for the major part of the variance of most tests designed to measure associational, expressional, and ideational fluency.

The Verbal Reasoning factor (Table 49) contains the factor loadings at all four grade levels. At grades nine and ten, reasoning tests have high loadings on this factor. For grades 11 and 12, only the Inference test (Rs-3) has appreciable loadings. However, the presence in this factor of tests that measured originality, fluency, sensitivity to problems, etc., in other studies makes its identification less certain. Nevertheless, the vocabulary test loadings for all grades firmly establish its verbal nature.

Flexibility of Closure (Table 50) contains high loadings for the tests purported to measure it but also it includes tests of length estimation, one measure of originality, and one measure of figural adaptive flexibility. In addition, at grade twelve all inductive reasoning tests have high loadings on this factor. The relationship between Flexibility of Closure and inductive reasoning has been previously reported (Tyler, 1965).

The Reasoning factor (Table 51) for grades ten, eleven, and twelve is defined by general reasoning tests, syllogistic reasoning tests and the figural adaptive flexibility tests for all three grades; however, two inference tests do not load on it at grade twelve. It should be noted that this factor does not appear for grade nine.

The Fluency in Making Verbal Transformations (Table 51a) for grades ten, eleven, and twelve is defined by Associations IV (Fa-3), Thing Categories (Fi-3), Gestalt Transformations (Re-1). This factor does not appear at grade nine.

The factors which are common for all grade levels have been described above. The remaining descriptions are

by grade level and deal with factors which emerged but were not common over all grades. Some are specific to grade level, some are unidentifiable, and some appear at two grade levels.

Grade Nine. Table 52 reveals that four of the factors for grade nine are not common over grade levels. Factor VIII is named Figural Reasoning because three of the four tests which define it involve figural content and some elements of reasoning appear in each of them. Factor X is defined by three tests each of which have small loadings on it. The factor cannot be interpreted so it was not named. Factor VII and Factor IX are specific. The former has a loading of .46 for Planning Air Maneuvers (X-23) and the latter has a loading of .33 for Nonsense Syllogisms (Rs-1). These factors have been named according to the KIT manual. Factor VII is named Figural Adaptive Flexibility and Factor IX is named Syllogistic Reasoning.

Grade Ten. Table 53 reveals that two factors for grade ten are not common over grades. Factor VI is named Semantic Redefinition. It is almost specific because it has only one large loading for Object Synthesis (Re-2) and one small loading for Logical Reasoning (Rs-2). Factor VII is named Deductive Reasoning. All three tests which have substantial loadings on it involve reasoning; Ship Destination (R-3) and Inferences (Rs-3) appear on reasoning factors in other studies and Hidden Figures (Cf-1) is associated with reasoning variables in several other factors in this study.

Grade Eleven. Table 54 reveals that two factors for grade eleven are not common for all grades. Factor III is named Deductive Reasoning. It is defined by tests which require reasoning from general principles or ideas to specific products. The factor is named Deductive Reasoning despite the loadings of two tests which are divergent production in nature, Plot Titles (O-2) and Planning Air Maneuvers (Xa-3). Factor X is named Sensitivity to Problems. It is a doublet and is primarily defined by the Apparatus test (Sep-1).

Grade Twelve. Table 55 reveals that two factors for grade twelve are not common over all grades. Factor X is named Syllogistic Reasoning. Factor VII is not named nor identified. It is a bipolar factor and the occurrence of such factors is highly unusual in the analyses of ability tests.

The factor scores are somewhat unexpected and disappointing. We had expected sixteen factors to emerge that

would be common for all grades. Instead only six common factors emerged and two others emerged for three of the four grade levels. Furthermore, the components of each factor are not identical at each grade level, but they are sufficiently similar to warrant giving them a common name. The majority of remaining factors are specific to grade level. These discrepancies from our expectancy seriously impede the interpretation of the regression analyses because of the differences between factors at the several grade levels. These discrepancies are undoubtedly attributable to any one or more of the following: factor differentiation within the students was different at the several age and grade levels, some of the tests were too difficult for the youngest groups who were tested, and some of the tests were too simple for the oldest groups with the consequence that these students were not differentiated by the tests.

Regression analyses. For each grade level, twenty-four regression equations were computed. These equations appear in Tables 56 through 59. To facilitate interpretation, the significant predictors for each process level for all forms for each grade were summarized and appear in Tables 60 and 60a through 63 and 63a. The first table in each set contains the identification of significant beta weights and the second table contains the numerical values of these significant beta weights and they are arranged according to common and specific factors.

It is important to remember that the factor analyses of the taxonomy-type tests showed the presence of content factors. Therefore, caution must be used when interpreting significant beta coefficients as indicating an aptitude-process relationship inasmuch as it might have arisen because of the particular demands of the content under way.

The results of the regression analyses are examined in several ways. Comparisons are made for process scores (a) by grade and form; (b) by grade over forms (c) by form over grades and (d) over grades and forms. Any conclusions based on the results of over-grade summaries must be held quite tentatively because the components of similarly named factors for different grades are not necessarily identical.

As an aid, summary predictor equations are used for some of the comparisons. Where used, the predictor components appear as the codes of cognitive ability factors according to the code which appears in Table 64.

TABLE 39

Unrotated Factor Matrix

Kit of Reference Tests

Grade 9

FACTORS

Variable	I	II	III	IV	V	VI	VII	VIII	IX	X
Cf 1	-427	068	-183	-051	-132	106	160	062	237	-024
Cf 2	-513	110	-381	033	116	051	045	038	036	250
Cf 3	-588	150	-407	068	193	082	-111	-057	-002	103
Cs 1	-176	061	-300	059	-445	-278	084	-132	-100	-075
Cs 2	-389	-016	-329	070	-431	-002	019	-153	-004	-034
Fa 1	-575	-045	225	-042	-049	-179	-195	032	104	-109
Fa 3	-431	-159	167	-041	223	-005	-075	-025	-063	014
Fe 2	-515	-051	140	-299	-051	-017	-196	041	117	070
Fe 3	-562	-021	074	-271	-065	212	-006	-217	046	004
Fi 1	-487	-124	074	-327	033	-001	057	-030	-185	068
Fi 3	-386	-057	-036	-324	-164	-156	-177	110	-049	072
I 1	-591	-049	041	-009	-023	-131	-041	-208	047	-080
I 2	-543	211	076	233	-052	058	023	-037	-144	-008
I 3	-383	074	-084	-157	022	075	041	-016	198	-003
Le 1	-348	172	-231	-056	120	-037	-012	144	244	-134
Le 2	-401	189	-326	061	023	-086	-265	074	-055	-025
Le 3	-398	152	-176	233	073	-104	-156	-018	201	-078
Ma 1	-276	-404	-128	158	-266	317	064	222	-019	191
Ma 3	-425	-415	107	175	-196	230	-008	126	010	-103
N 1	-545	-544	-099	109	264	-265	117	009	-028	012
N 2	-573	-498	-104	116	072	007	051	-003	013	049
N 3	-450	-576	-069	099	234	-226	173	-129	040	014
O 2	-632	067	010	-176	055	104	-061	-181	045	-131
R 3	-506	429	187	095	100	025	149	086	033	-027
R 4	-471	119	137	248	042	-124	-000	185	003	-064
Re 1	-361	357	001	007	-120	-216	159	061	048	080
Re 2	-327	148	-167	-353	-101	-176	255	170	-042	-033
Rs 1	-142	022	164	065	037	-168	-069	-134	200	124
Rs 2	-581	144	249	277	-022	045	-073	-081	-020	098
Rs 3	-595	235	357	127	086	-038	054	-033	-061	-044
Sep1	-332	-131	000	-435	163	079	032	033	-184	-046
Sep2	-385	-044	065	-009	104	066	108	042	-031	-272
V 1	-616	019	348	-007	-106	-047	-014	115	006	064
V 3	-345	058	289	-005	-035	025	052	109	-085	176
Xa 1	-352	222	-285	011	235	194	-028	102	-122	-060
Xa 3	-244	327	-013	-034	022	118	295	-160	-031	104

TABLE 38

Unrotated Factor Matrix

Kit of Reference Tests

Grade 10

FACTORS

Variable	I	II	III	IV	V	VI	VII	VIII	IX	X
Cf 1	-388	-342	-170	295	065	072	290	-103	000	003
Cf 2	-494	181	-008	263	116	169	-168	023	-142	-053
Cf 3	-519	-283	060	362	035	296	-112	009	014	-049
Cs 1	-287	-113	276	243	273	-181	-110	-008	055	370
Cs 2	-352	010	046	231	184	-109	-088	150	193	106
Fa 1	-562	116	049	-124	163	-098	-073	145	090	130
Fa 3	-361	-185	271	-125	003	-042	094	-038	-164	034
Fe 2	-445	248	154	-088	048	171	-045	-211	104	017
Fe 3	-561	165	112	-209	029	180	-015	130	107	091
Fi 1	-353	038	500	-072	-168	-043	133	015	051	-046
Fi 3	-420	025	503	-000	171	-118	187	-143	-236	056
I 1	-624	011	004	099	-237	-057	056	025	-032	202
I 2	-526	-102	-095	132	-192	-121	-026	032	054	096
I 3	-367	-144	026	065	-236	-067	-091	-004	-028	-024
Le 1	-391	128	-116	074	217	110	-060	-044	177	-138
Le 2	-468	-199	024	181	035	117	-038	-067	-048	-154
Le 3	-468	-153	-090	035	-083	099	-165	157	085	004
Ma 1	-233	188	-133	049	115	258	307	255	-234	-057
Ma 3	-477	361	-106	-082	-034	077	210	377	029	-055
N 1	-437	637	-067	314	008	-030	-059	-106	-068	-058
N 2	-540	572	-164	125	-085	-151	-052	-075	-099	-094
N 3	-411	720	020	194	-023	-119	014	-176	035	-048
O 1	-142	012	476	-012	-027	185	101	035	159	067
O 2	-487	-156	143	-100	109	093	-037	-037	139	-184
R 3	-546	-369	-257	-066	-029	-117	242	-156	108	-034
R 4	-675	144	-181	-150	-076	-125	044	018	-059	166
Re 1	-478	-313	-089	-120	259	-114	-052	-015	-252	-075
Re 2	-392	-227	474	087	055	-412	023	151	119	-307
Rs 1	-259	-110	-036	-141	-101	-008	-329	-100	-037	-059
Rs 2	-664	-061	-123	-136	-211	-226	-097	154	-016	-088
Rs 3	-460	-173	-392	-021	047	-064	366	-166	220	019
Sep1	269	089	401	-215	-281	267	036	-147	120	-012
Sep2	-418	060	058	-293	035	205	-007	-059	-090	-029
V 1	607	014	-201	-387	203	087	-110	-139	-045	038
V 3	-593	-055	-194	-234	161	-109	-113	059	002	020
Xa 1	-496	-244	-014	108	-285	097	-035	-013	-145	041
Xa 3	-222	-223	-102	081	-323	067	-032	-004	-026	066

TABLE 39

Unrotated Factor Matrix

Kit of Reference Tests

Grade 11

FACTORS

Variable		I	II	III	IV	V	VI	VII	VIII	IX	X
Cf	1	-594	-101	-024	-007	-107	-118	046	077	177	103
Cf	2	-600	-200	-172	-214	-183	-102	207	-308	175	143
Cf	3	-648	-316	-201	-064	-319	-083	-066	-127	053	-054
Cs	1	-238	-306	-350	-575	259	222	058	-014	-104	-085
Cs	2	-442	-138	-315	-477	351	033	046	012	-157	-100
Fa	1	-543	117	-091	081	332	-119	-031	-092	-011	167
Fa	3	-574	120	-186	-038	457	-046	050	167	208	-007
Fe	2	-277	074	163	150	-072	166	163	130	-191	019
Fe	3	-482	211	-106	159	-002	091	133	-073	-116	122
Fi	1	-513	082	-335	311	-001	186	-172	-087	-165	046
Fi	3	-540	113	-271	140	247	032	-148	-039	-030	-213
I	1	-605	022	036	-181	054	028	263	098	064	-160
I	2	-428	-162	208	141	043	090	234	103	-037	092
I	3	-460	-266	-061	156	009	150	-137	154	-056	-163
Le	1	-388	-175	-194	-124	-329	-131	049	018	006	-105
Le	2	-490	-234	-137	-030	-274	-133	-179	-023	023	001
Le	3	-409	-192	-006	-010	-260	-051	-050	-140	-238	135
Ma	1	-443	330	-078	-030	-095	-346	204	038	147	008
Ma	3	-425	457	-065	014	-032	-220	302	051	-155	-132
N	1	-464	538	028	-343	-195	155	-181	-025	010	-015
N	2	-517	399	208	-261	-13[illegible]	009	-247	030	-006	026
N	3	-475	602	004	-188	-223	298	-050	-058	118	036
O	1	-255	-004	-128	259	-014	475	170	004	219	021
O	2	-599	-185	-111	104	-039	-086	-076	-091	-039	-137
R	3	-566	-275	308	-017	-047	066	-067	113	-048	059
R	4	-673	004	198	-111	-039	-016	-032	301	-089	104
Re	1	-524	-158	114	037	180	-189	-260	157	217	120
Re	2	-331	-123	-359	155	136	-074	-050	098	098	-011
Rs	1	-184	-085	295	-121	128	243	-030	-102	199	-055
Rs	2	-598	-028	327	047	-071	058	019	140	-167	039
Rs	3	-608	020	406	077	079	-046	-040	067	-012	-185
Sep	1	-236	219	-319	469	-128	-016	-140	037	086	-320
Sep	2	-484	021	-016	163	021	137	117	-154	-010	017
V	1	-554	057	279	172	330	-144	050	-273	-110	-018
V	3	-526	026	385	063	214	-057	-048	-305	-032	-1[illegible]3
Xa	1	-468	-250	075	096	-173	031	148	171	-034	019
Xa	3	-273	-360	263	077	-082	200	048	-120	125	-040

TABLE 40

Unrotated Factor Matrix

KIT Cf Reference Tests

Grade 12

FACTORS

Variable	I	II	III	IV	V	VI	VII	VIII	IX	X
Cf 1	594	038	009	216	-021	-028	-213	-102	-091	-070
Cf 2	552	-043	032	355	-053	064	022	-219	-132	109
Cf 3	068	053	074	326	-163	084	-034	-069	-134	063
Cs 1	246	044	571	181	062	-192	303	253	-011	015
Cs 2	394	054	580	192	075	-007	071	019	-007	061
Fa 1	486	035	110	-306	097	025	101	006	-158	-125
Fa 3	417	-083	498	-255	080	096	-083	-173	089	-101
Fe 2	321	001	-169	-145	-154	023	180	053	-093	-053
Fe 3	505	131	-216	-184	-032	-095	218	-105	008	110
Fi 1	371	057	-128	-269	-356	148	089	078	088	198
Fi 3	505	058	342	-195	-161	-020	-114	144	-164	090
I 1	620	029	-005	152	-006	020	107	-150	039	-147
I 2	451	-195	-128	183	-014	-070	259	-085	-017	-010
I 3	445	-234	-070	217	-101	136	049	060	-005	-087
Le 1	478	088	-057	173	-097	-090	-048	-084	-138	045
Le 2	558	-143	044	145	-235	-049	-188	-049	-010	121
Le 3	452	-100	-113	123	-179	-179	-224	-020	-016	041
Ma 1	308	275	094	-176	-062	-437	-112	-143	207	-090
Ma 3	504	319	089	-137	-140	-281	058	-207	195	-099
N 1	378	747	-076	124	181	144	091	087	047	050
N 2	582	533	-118	008	217	121	-155	076	005	-004
N 3	427	717	-156	039	155	173	-082	058	084	028
O 1	300	132	-001	-271	-301	-139	009	144	-127	107
O 2	548	-237	080	-020	-045	285	065	-011	-026	-124
R 3	579	-233	-004	044	161	-073	-118	274	018	-177
R 4	676	-030	-245	056	127	-034	-050	167	-020	-073
Re 1	429	-274	125	-099	101	110	-344	101	071	117
Re 2	361	-121	172	-179	-215	208	064	082	337	145
Rs 1	105	-188	-206	-145	154	-228	-002	021	-121	220
Rs 2	609	-133	-227	-062	189	-193	103	025	-060	-022
Rs 3	630	-195	-145	-219	213	-195	016	086	067	059
Sep 1	251	003	-149	-418	-315	212	-079	-057	-085	-226
Sep 2	347	063	022	-146	-113	-008	124	124	-191	-118
V 1	498	-243	060	-268	398	106	-052	-176	-048	109
V 3	464	-267	-046	-197	298	176	169	-151	052	077
Xa 1	470	-200	-216	322	-141	076	134	040	242	024
Xa 3	312	-348	-145	141	011	-071	-076	150	224	-122

TABLE 41

Rotated Factor Matrix

Kit Of Reference Tests

Grade 9

FACTORS

Variable	I	II	III	IV	V	VI	VII	VIII	IX	X
Cf 1	-161	-070	-186	-150	-177	239	147	388	039	-087
Cf 2	-099	-145	-580	-164	-079	133	197	175	116	055
Cf 3	-141	-168	-680	-140	-081	069	154	162	059	-119
Cs 1	-014	-031	-078	-021	-660	008	008	033	-017	026
Cs 2	-041	-045	-216	-098	-555	264	072	096	042	-147
Fa 1	-431	-176	-080	-340	-084	043	-235	151	167	-164
Fa 3	-278	-122	015	-316	-163	237	-060	-043	075	-121
F[illegible]	-222	-057	-105	-531	011	106	-077	158	192	-115
F[illegible]	-212	-078	-124	-416	-054	173	200	103	090	-371
Fi 1	-178	-218	-080	-537	-024	053	140	-005	-075	-080
Fi 3	-086	-038	-128	-521	-176	055	-123	079	065	062
I 1	-288	-203	-208	-245	-062	143	089	094	082	-384
I 2	-528	-046	-274	-051	-148	113	121	000	-027	-136
I 3	-106	-048	-171	-227	-014	072	110	316	070	-121
Le 1	-097	-053	-314	-084	-027	[illegible]	-040	444	-000	-052
Le 2	-135	-005	-545	-120	-186	[illegible]	-132	107	013	-012
Le 3	-291	-119	-451	008	-179	[illegible]	-082	-052	-065	-067
M 1	-009	-172	-059	-063	-080	679	002	023	-087	031
M 2	-226	-263	025	-117	-040	568	-137	031	-057	-177
N 1	-140	-812	-161	-188	-029	136	-074	054	021	005
N 2	-132	-598	-185	-202	-050	376	-013	051	052	-105
N 3	-059	-804	-049	-132	-037	127	001	036	086	-082
O 2	-279	-127	-256	-375	-060	031	092	184	032	-395
R 3	-602	056	-197	-068	028	-076	199	257	-029	-061
R 4	-529	-147	-181	-012	-041	057	-100	149	000	031
Re 1	-375	049	-128	-106	-252	-122	163	240	070	131
Re 2	-084	-044	-072	-376	-243	-099	157	331	-192	154
Rs 1	-169	-093	026	-019	005	-084	-010	036	334	-036
Rs 2	-609	-081	-211	-071	-040	152	075	-035	189	-162
Rs 3	-692	-098	-109	-159	018	-044	095	064	034	-168
Sep1	-004	-168	-090	-521	116	-013	086	043	-213	-106
Sep2	-260	-203	-065	-126	022	054	-016	177	-211	-223
V 1	-582	-109	-014	-353	-031	175	-024	097	113	-039
V 2	-410	-020	011	-230	062	122	109	-027	051	076
Xa 1	-122	-005	-513	-070	072	016	117	145	-219	-081
Xa 2	-219	069	-111	-037	-072	-077	457	111	-031	-082

TABLE 42

Rotated Factor Matrix

Kit of Reference Tests

Grade 10

FACTORS

Variable	I	II	III	IV	V	VI	VII	VIII	IX	X
Cf 1	-239	-033	-033	029	369	-005	499	114	-142	095
Cf 2	-261	086	009	-180	540	-029	016	069	-131	130
Cf 3	-323	016	140	-065	650	-017	107	047	-006	151
Cs 1	-055	045	089	-051	164	-100	024	-115	-159	631
Cs 2	-066	126	064	-106	326	-054	047	111	165	342
Fa 1	-126	217	135	-452	064	-159	059	139	028	308
Fa 3	-171	-070	297	-396	070	-151	077	034	-300	088
Fe 2	-009	341	376	-288	154	072	071	-040	023	058
Fe 3	-141	187	357	-437	102	-014	027	221	116	147
Fi 1	-157	103	546	-035	-008	-290	003	038	-091	076
Fi 3	008	163	438	-113	095	-214	063	046	-489	237
I 1	-517	258	191	-179	071	-032	174	126	-069	211
I 2	-479	171	026	-163	122	-112	195	025	015	164
I 3	-411	067	088	-103	121	-129	047	-054	-040	010
Le 1	049	275	019	-264	319	-048	161	048	144	049
Le 2	-236	075	100	-143	448	-108	151	012	-107	007
Le 3	-366	018	038	-276	270	-072	043	080	144	093
M 1	019	141	021	-069	149	091	071	577	-096	-064
M 3	-136	339	124	-265	-011	-118	066	556	163	015
N 1	-062	817	025	-053	154	028	-036	130	-017	072
N 2	-195	777	-014	-229	026	-054	021	-136	-039	-019
N 3	008	861	142	-065	-012	-026	018	065	013	066
O 1	009	-039	497	056	086	-055	-088	083	-065	132
O 2	-087	014	268	-339	317	-218	159	-021	021	009
R 3	-340	-041	-004	-303	140	-127	603	-005	-081	-005
R 4	-374	336	070	-471	-048	-012	214	178	-077	142
Re 1	-153	-082	-096	-459	262	-171	159	040	-348	063
Re 2	-124	005	236	-053	160	-768	057	-071	-131	165
Rs 1	-246	023	-006	-321	114	-013	-088	-213	-001	-082
Re 2	-517	190	018	-445	034	-303	110	095	000	000
Rs 3	-182	087	-065	-238	076	-001	704	090	045	017
Sep1	-141	070	648	-100	008	032	-029	-057	067	-107
Sep2	-075	087	291	-427	105	067	034	110	-086	-033
V 1	-097	145	086	-732	130	097	190	027	-085	006
V 3	-200	118	-046	-609	108	-122	166	077	-033	120
Xa 1	-554	016	142	-123	248	006	103	045	-125	001
Xa 3	-451	-064	029	001	092	063	091	-018	023	-042

TABLE 43

Rotated Factor Matrix

Kit Of Reference Tests

Grade 11

FACTORS

Variable	I	II	III	IV	V	VI	VII	VIII	IX	X
Cf 1	-306	129	124	-018	055	073	195	-414	306	008
Cf 2	-042	103	234	-182	156	105	229	-666	155	228
Cf 3	-200	091	089	-141	087	122	037	-756	120	-118
Cs 1	-003	008	090	-825	-056	022	-060	-184	041	096
Cs 2	-070	115	-084	-756	111	079	050	-179	230	003
Fa 1	-103	116	-032	-127	325	285	152	-092	471	055
Fa 3	-169	127	139	-331	157	145	297	016	586	-112
Fe 2	-170	043	029	-065	-087	419	141	-061	-023	-092
Fe 3	-127	197	063	001	154	465	203	-128	130	016
Fi 1	-049	152	014	-045	119	601	-078	-239	232	-240
Fi 3	-062	125	001	-072	174	483	007	-125	508	-072
I 1	-352	190	198	-295	164	070	378	-298	098	-048
I 2	-467	-069	188	-000	154	191	123	-101	077	092
I 3	-371	-031	097	-158	077	191	-121	-240	142	-312
Le 1	-134	068	-010	-120	-075	032	166	-538	-016	-111
Le 2	-190	103	-043	-045	026	059	-033	-589	158	-116
Le 3	-218	069	-118	-011	124	221	-066	-472	-043	105
M 1	-050	254	-052	081	078	053	562	-221	235	-010
M 3	-097	265	-151	-018	160	256	612	-039	015	-090
N 1	-060	811	-007	-110	067	099	139	-109	024	-025
N 2	-287	727	-082	-040	210	033	109	-174	157	-009
N 3	-031	802	201	001	036	239	190	-063	022	-019
O 1	-094	041	536	011	-063	362	-007	-012	068	-120
O 2	-220	024	022	-124	261	192	060	-452	186	-217
R 3	-594	096	089	-062	205	032	-091	-283	117	042
R 4	-614	304	-054	-112	078	109	136	-209	225	029
Re 1	-359	083	007	-017	175	-092	-040	-220	555	-052
Re 2	-026	-112	037	-014	-035	206	077	-211	384	-027
Rs 1	-180	141	334	-100	234	-144	-128	033	026	060
Rs 2	-599	212	007	011	235	171	069	-170	046	016
Rs 3	-531	203	057	011	439	-009	133	-102	129	-155
Sep1	106	069	035	202	019	299	136	-125	140	-610
Sep2	-161	096	210	-025	254	337	107	-191	101	-026
V 1	-242	043	001	-010	678	185	154	-056	230	049
V 3	-256	156	086	-030	700	010	061	-103	090	-055
Xa 1	-480	040	124	-014	011	158	110	-334	045	-035
Xa 3	-311	085	353	020	207	030	-159	-241	-059	023

TABLE 44

Rotated Factor Matrix

Kit of Reference Tests

Grade 12

FACTORS

Variable	I	II	III	IV	V	VI	VII	VIII	IX	X
Cf 1	561	203	045	-155	-094	-153	-138	185	-070	-017
Cf 2	680	121	124	-171	-017	009	072	033	037	029
Cf 3	706	219	185	-102	-168	-021	-037	103	092	-035
Cs 1	076	017	772	-019	-046	-074	046	073	013	-031
Cs 2	285	094	606	-222	030	-094	-104	-019	059	-128
Fa 1	087	150	148	-397	-412	-122	-012	082	-040	049
Fa 3	105	-032	273	-552	-121	-256	-201	-000	111	-208
Fe 2	137	068	-045	-060	-386	-017	143	106	091	079
Fe 3	221	239	-042	-184	-311	-218	226	045	167	262
Fi 1	140	113	-091	-037	-408	-052	-008	[illegible]	474	077
Fi 3	233	087	324	-186	-392	-138	-339	-003	141	-033
I 1	470	197	098	-245	-154	-174	175	215	030	-076
I 2	406	-023	085	-143	-110	-017	302	233	049	140
I 3	418	-013	044	-114	-133	123	069	328	101	-059
Le 1	483	174	053	-028	-152	-122	-013	060	-019	100
Le 2	562	-016	060	-084	-128	-132	-189	166	193	067
Le 3	457	003	-057	018	-111	-191	-161	224	060	142
M 1	065	149	074	-018	-095	-664	-061	039	007	064
M 3	224	239	111	-102	-223	-610	096	019	114	-021
N 1	121	855	134	028	-080	-093	109	-066	027	-046
N 2	234	760	012	-167	-154	-144	-113	113	-012	026
N 3	139	859	-028	-021	-101	-144	-018	-014	054	-052
O 1	088	067	066	073	-457	-180	-133	-036	174	163
O 2	346	003	078	-394	-266	103	015	248	152	-145
R 3	250	085	163	-242	-166	-036	-149	572	-047	097
R 4	364	307	-003	-194	-231	-050	-009	465	-004	198
Re 1	210	-004	039	-360	-014	009	-406	275	180	086
Re 2	087	-006	121	-213	-150	072	-058	142	562	-099
Rs 1	014	-076	-073	-101	-038	007	-017	069	-033	474
Rs 2	289	141	024	-289	-228	-138	132	363	-051	352
Rs 3	179	108	046	-382	-211	-200	-007	406	094	393
Sep1	036	-008	-314	-167	-534	-087	[illegible]	[illegible]	121	-175
Sep2	123	096	126	-080	-431	-039	016	083	-021	018
V 1	150	068	020	-702	-055	-023	-085	117	025	238
V 3	141	047	-002	-610	-081	050	156	173	143	178
Xa 1	457	052	-013	-005	008	039	229	408	293	011
Xa 3	202	-110	-033	-068	040	-032	-001	525	100	047

TABLE 45

Factor Loadings: Number Facility

Name of Test	Grade: 9	10	11	12
	Factor: II	II	II	II
	Factor Loadings			
Addition	.812	.818	.811	.855
Division	.598	.777	.727	.760
Subtraction and Multiplication	.804	.861	.802	.859

TABLE 46

Factor Loadings: Associative Memory

Name of Test	Grade: 9	10	11	12
	Factor: VI	VIII	VII	VI
	Factor Loadings			
Picture-Number	.679	.577	.562	.664
First and Last Name	.568	.556	.612	.610

TABLE 47

Factor Loadings: Speed of Closure

Name of Test	Factor Loadings			
Grade:	9	10	11	12
Factor:	V	X	IV	III
Gestalt Completion	.660	.631	.825	.772
Concealed Words	.555	.342	.756	.606

TABLE 48

Factor Loadings: Fluency

Name of Test				
Grade:	9	10	11	12
Factor:	IV	III	VI	V
Controlled Associations (a1)	.340			.412
Associations IV (fa3)	.316	.297		
Simile Interpretation (fe2)	.513	.376	.419	.386
Word Arrangements (fe3)	.416	.357	.465	.311
Topics (fi1)	.537	.546	.601	.408
Thing Categories (fi3)	.521	.438	.483	.392
Plot Titles (01)	*	.497	.362	.457
Symbol Production (02)	.375			
Apparatus Test (Sep1)	.521	.648	.299	.534
Seeing Problems (Sep2)		.291	.337	.431

*Not in Ninth Grade Battery

TABLE 49

Factor Loadings: Verbal Reasoning

Name of Test	Grade: 9 Factor: I	10 IV	11 V	12 IV
	Factor Loadings			
Vocabulary (V1)	.582	.732	.678	.702
Wide Range Vocabulary	.410	.609	.700	.610
Controlled Associations (Fa1)	.431	.452	.325	.397
Associations IV (Fa3)				.552
Symbol Production (02)		.339		.394
Ship Destination (R3)	.602	.303		
Necessary Arithmetic Operations (R4)	.529	.471		
Gestalt Transformation (Re1)	.375	.459		.360
Object Synthesis (Re2)				
Nonsense Syllogisms (Rs1)		.321		
Logical Reasoning (Rs2)	.609	.445		
Inference (Rs3)	.692		.439	.382
Seeing Problems (Sep2)		.427		
Locations (I2)	.528			
Word Arrangements (Fe3)		.437		

TABLE 50

Factor Loadings: Flexibility of Closure

Name of Test	Grade: 9 Factor: III	10 V	11 VIII	12 I
Hidden Figures (Cf1)		.369	.414	.561
Hidden Patterns (Cf2)	.580	.540	.666	.680
Copying (Cf3)	.680	.650	.756	.706
Estimation of Length (Le1)	.314	.319	.538	.483
Shorted Road (Le2)	.545	.448	.589	.562
Nearer Point (Le3)	.451		.472	.457
Letter Sets (I1)				.470
Locations (I2)				.406
Figure Classification (I3)				.418
Symbol Production (O2)		.317	.452	.346
Match Problems II (Xa1)	.513		.334	.457
Concealed Words (C52)		.326		

TABLE 51

Factor Loadings: Reasoning

Name of Test	Grade: 10 Factor: I	11 I	12 VIII
Hidden Figures (Cf1)		.306	
Copying (Cf3)	.323		
Letter Sets (I1)	.517	.352	
Locations (I2)	.479	.467	
Figure Classification (I3)	.411	.371	.328
Nearer Point (Le3)	.366		
Ship Destination (R3)	.340	.594	.572
Necessary Arithmetic Operations (R1)	.374	.614	.465
Gestalt Transformation (Re1)		.359	
Logical Reasoning (Rs2)	.517	.599	.363
Inference (Rs3)		.531	.406
Match Problems II (Xa1)	.544	.480	.408
Planning Air Manuevers (Xa3)	.451	.311	.535

TABLE 51a

Factor Loadings: Fluency in Making Verbal Transformations

Name of Test	Grade: 10 Factor: IX	11 IX	12 IX
Controlled Associations (Fa1)		.471	
Associations IV (Fa3)	.300	.586	
Topics (Fi1)			.474
Thing Categories (Fi3)	.489	.506	
Gestalt Transformation (Re1)	.348	.555	
Object Synthesis (Re2)		.384	.562

TABLE 52

Common Factors Appearing Only at Grade 9

Figural Reasoning: Factor VIII

Name of Test	Loadings
Hidden Figures (Cf-1)	.388
Figure Classification (I-3)	.316
Estimation of Length (Le-1)	.444
Object Synthesis (Re-2)	.331

Unidentified: Factor X

Name of Test	Loadings
Word Arrangements (Fe-3)	.371
Letter Sets (I-1)	.384
Symbol Production (O-2)	.395

TABLE 53

Common Factors Appearing only at Grade 10

Semantic Redefinition: Factor VI

Name of Test	Factor Loadings
Object Synthesis (Re-2)	.768
Logical Reasoning (Rs-2)	.303

Deductive Reasoning: Factor VII

Name of Test	Factor Loadings
Hidden Figures (Cf-1)	.499
Ship Destination (R-3)	.603
Inference (Rs-3)	.704

Fluency in Making Verbal Transformations: Factor IX

Name of Test	Factor Loadings
Associations IV (Fa-3)	.300
Thing Categories (Fi-3)	.489
Gestalt Transformations (Re-1)	.348

TABLE 54

Common Factors Appearing only at Grade 11

Deductive Reasoning: Factor III

Name of Test	Factor Loadings
Plot Titles (O-1)	.536
Nonsense Syllogisms (Rs-1)	.334
Planning Air Manuevers (Xa-3)	.353

Fluency in Making Verbal Transformations: Factor IX

Name of Test	Factor Loadings
Controlled Associations (Fa-1)	.471
Associations IV (Fa-3)	.586
Thing Categories (Fi-3)	.506
Gestalt Transformation (Re-1)	.555
Object Synthesis (Re-2)	.384

Sensitivity to Problems: Factor X

Name of Test	Factor Loadings
Figure Classification (I-3)	.312
Apparatus Test (Sep-1)	.610

TABLE 55

Common Factors Appearing only at Grade 12

Unidentified: Factor VII

Name of Test	Factor Loadings
Thing Categories (Fi-2)	-.334
Locations (I-2)	.302
Gestalt Transformations (Re-1)	-.406

Fluency in Making Verbal Transformations: Factor IX

Name of Test	Factor Loadings
Topics (Fi-1)	.474
Object Synthesis (Re-2)	.562

Deductive Reasoning: Factor X

Name of Test	Factor Loadings
Nonsense Syllogisms (Rs-1)	.474
Logical Reasoning (Rs-2)	.352
Inference (Rs-3)	.393

TABLE 56

Regression Analyses Grade 9

N = 300

Score Predicted	Verbal Reasoning B_1	Number Facility B_2	Flexibility of Closure B_3	Fluency B_4	Speed of Closure B_5	Memory B_6	Figural Adaptive Flexibility B_7	Figural Reasoning B_8	Syllogistic Reasoning B_9	Uninterpreted B_{10}	M	Multiple Correlation Coefficient
K	-.428	-.114	.015	-.039	.032	.177	.086	-.008	.081	-.067	62.9	.568
C	-.352	-.044	-.102	-.187	-.087	.065	-.004	.029	-.078	.063	84.9	.477
Ap	-.424	-.004	-.101	-.157	-.033	.001	.066	-.049	-.001	-.060	87.9	.525
An	-.284	-.048	-.132	-.076	-.049	.011	.030	.013	.011	-.011	76.6	.379
Sy	-.155	-.112	-.019	-.074	.038	.006	.002	.008	.121	-.008	59.5	.292
Ev	-.294	-.155	-.030	-.009	.102	-.013	.026	-.016	-.010	-.121	72.7	.419
K	-.429	-.090	-.118	-.073	.081	.181	.054	-.082	.121	-.128	73.7	.591
C	-.422	-.028	-.247	-.114	-.092	.050	.030	-.037	-.02[illegible]	-.026	95.7	.550
Ap	-.552	.028	-.117	-.075	-.034	.034	.057	-.001	-.106	.030	87.2	.593
An	-.413	.027	-.207	-.079	-.094	.038	.020	.137	-.051	.016	79.9	.546
Sy	-.215	.012	-.039	-.162	-.205	-.122	.063	.042	.082	.117	71.1	.401
Ev	-.375	.185	.025	-.083	.043	.152	.024	-.005	.006	-.151	74.3	.541
K	-.432	-.028	.035	-.099	.033	.239	.032	-.007	.098	-.038	58.2	.574
C	-.468	-.169	-.002	-.181	-.026	.115	.096	.012	.078	.045	74.9	.612
Ap	-.477	-.101	-.055	-.160	-.022	.078	.101	-.049	-.045	.029	85.0	.574
An	-.419	-.118	-.047	-.139	.048	.098	.066	-.077	-.010	-.026	81.2	.526
Sy	-.388	-.194	-.020	-.166	.039	.079	.012	.059	.089	-.159	82.1	.598
Ev	-.257	-.192	-.081	-.186	.021	.112	.067	.034	.115	.096	63.1	.484
K	-.389	-.139	-.013	-.108	.067	.226	-.002	-.005	.020	-.115	72.9	.587
C	-.509	-.055	-.031	-.197	-.097	.149	.098	-.069	-.001	.039	83.6	.628
Ap	-.463	-.023	-.064	-.194	-.015	.081	.212	-.046	.093	-.006	72.9	.617
An	-.349	.012	.042	-.150	-.083	.192	.044	.087	-.026	.053	58.6	.472
Sy	-.359	-.086	-.063	-.097	-.014	.176	.021	-.041	.005	-.011	73.2	.475
Ev	-.415	-.102	-.049	-.142	-.026	.032	.031	-.020	.030	-.021	82.2	.509

TABLE 57

Regression Analyses Grade 10
N = 275

Score Predicted	Reasoning B_1	Number Facility B_2	Fluency B_3	Verbal Reasoning B_4	Flexibility of Closure B_5	Semantic Redefinition B_6	Deductive Reasoning B_7	Memory B_8	Fluency in Making Verbal Transformations B_9	Speed of Closure B_{10}	M	Multiple Correlation Coefficient
K	-.216	.265	-.001	-.362	.066	.012	.243	.063	.075	.082	38.13	.629
C	-.188	.269	-.035	-.147	.073	-.095	.228	.083	-.106	.012	44.96	.505
Ap	-.250	.170	.013	-.267	.093	-.067	.234	.001	-.091	-.010	58.40	.546
An	-.150	.143	-.013	-.190	.084	-.059	.147	.044	.002	.097	44.50	.402
Sy	-.091	.163	-.052	-.242	-.010	-.020	.134	-.003	.065	.086	47.50	.390
Ev	-.172	.188	.004	-.184	.062	-.057	.158	.116	-.082	.112	34.00	.460
K	-.214	.232	.023	-.342	-.007	.016	.339	-.094	.072	.156	31.00	.617
C	-.273	.239	-.027	-.299	.097	-.043	.314	.032	-.072	-.001	51.20	.601
Ap	-.254	.093	-.086	-.346	.161	-.113	.426	.003	-.116	.017	51.70	.666
An	-.202	.097	.031	-.290	.053	-.026	.186	.024	-.026	.057	54.50	.455
Sy	-.161	-.033	.125	-.242	.129	-.117	.146	-.132	-.045	.025	68.80	.459
Ev	-.143	.221	.063	-.317	-.052	-.019	.182	.015	.160	.064	40.70	.526
K	-.212	.223	.021	-.325	-.002	.035	.362	.131	.045	-.002	37.80	.592
C	-.244	.288	.030	-.339	[illegible]	-.037	.349	.058	-.053	.013	45.50	.640
Ap	-.241	.232	-.029	-.361	-.017	.009	.355	.139	-.057	.020	46.90	.629
An	-.161	.190	-.029	.355	.055	-.027	.256	.076	.047	.057	44.00	.537
Sy	-.180	.169	.101	-.344	.002	-.006	.182	.094	.041	.025	45.40	.535
Ev	.000	.235	.126	-.263	-.051	.050	.131	.071	.018	.045	31.90	.430
K	-.136	.252	-.031	-.292	-.043	-.008	.229	.227	.116	.100	28.60	.582
C	-.214	.128	-.010	-.345	.016	-.079	.279	.046	.013	.031	54.80	.544
Ap	-.206	.178	-.015	-.330	.011	-.058	.281	.051	-.026	-.003	54.40	.525
An	-.186	.110	.027	-.166	.080	-.079	.287	.062	-.079	.002	46.50	.443
Sy	-.108	.139	-.027	-.242	[illegible]	-.023	.196	-.037	.032	-.003	49.70	.372
Ev	-.205	.206	-.010	-.212	.109	.015	.191	.028	.033	-.038	48.30	.416

TABLE 58

Regression Analyses Grade 11
N = 282

Score Predicted	Reasoning B_1	Number Facility B_2	Deductive Reasoning B_3	Speed of Closure B_4	Verbal Reasoning B_5	Fluency B_6	Memory B_7	Flexibility of Closure B_8	Fluency in making Verbal Transformations B_9	Sensitivity to Problems B_{10}	M	Multiple Correlation Coefficient
K	-.226	.089	.044	-.036	.337	.105	.250	.035	-.032	-.045	23.30	.576
C	-.215	.097	.020	-.098	.155	.109	.121	-.097	.107	-.041	31.70	.459
Ap	-.252	.122	.084	.008	.225	.105	.161	-.079	.135	-.048	26.30	.552
An	-.199	.109	.080	.041	.140	.131	.083	-.098	.130	-.067	32.00	.460
Sy	-.130	.127	.040	.062	.189	.085	-.011	-.012	-.013	.024	33.10	.365
Ev	-.184	.071	.041	-.007	.121	.169	.126	.064	-.100	.005	34.20	.381
K	-.307	.174	-.010	-.046	.192	.084	.216	-.038	-.007	-.045	38.80	.532
C	-.336	[illegible]	.114	-.057	.249	-.004	.055	-.197	.079	-.045	49.00	.605
Ap	-.375	.134	.108	-.064	.232	.042	.082	-.115	.186	-.019	38.90	.628
An	-.277	.120	.064	-.011	.177	.019	-.001	-.212	.081	-.031	53.20	.500
Sy	-.205	.088	.093	.056	.194	-.022	-.112	-.231	.121	-.065	53.30	.502
Ev	-.172	.132	.145	-.024	.148	.170	.166	-.042	-.026	.013	23.80	.448
K	-.309	.140	-.005	.073	.237	.043	.119	-.022	.068	.061	29.10	.527
C	-.320	.162	-.060	.025	.275	.059	-.001	-.033	.076	-.001	40.40	.557
Ap	-.336	.132	.014	-.005	.295	-.033	.053	-.113	.097	-.061	47.30	.606
An	-.313	.108	-.024	.053	.266	.015	.078	-.044	.064	-.056	42.60	.524
Sy	-.220	.142	.006	.037	.344	.144	.104	-.073	-.083	.009	29.00	.548
Ev	-.213	.148	.006	.004	.131	.105	.093	-.045	-.050	-.033	42.10	.378
K	-.234	.209	.012	.067	.219	.145	.231	.034	-.011	.026	14.30	.538
C	-.313	.146	.042	-.009	.318	.137	.122	-.042	.055	-.047	29.00	.602
Ap	-.320	.185	.063	.015	.279	.112	.149	-.104	.091	.056	23.10	.616
An	-.243	.144	.075	-.006	.171	.111	.059	-.044	.060	-.013	33.70	.437
Sy	-.167	.153	.076	.034	.231	.062	.021	-.040	-.014	-.009	32.00	.402
Ev	-.189	.159	.138	.099	.225	.058	.118	-.005	-.018	.045	17.50	.471

TABLE 59

Regression Analyses Grade 12
N = 267

Score Predicted	Flexibility of Closure B_1	Number Facility B_2	Speed of Closure B_3	Verbal Reasoning B_4	Fluency B_5	Memory B_6	Uninterpreted B_7	Reasoning B_8	Fluency in Making Verbal Transformations B_9	Syllogistic Reasoning B_{10}	M	Multiple Correlation Coefficient
K	.219	.030	-.095	-.113	.065	-.003	.152	.127	.066	.129	20.50	.433
C	.157	.095	-.030	-.184	-.069	.003	.085	.148	.037	.005	37.20	.396
Ap	.240	.131	-.042	-.133	-.077	.052	.021	.194	.064	.139	19.90	.509
An	.201	.093	.075	-.108	.606	.003	-.035	.114	.086	.059	24.60	.374
Sy	.092	.053	-.045	-.064	-.057	.002	.081	.232	.069	.168	22.60	.428
Ev	-.004	.144	-.176	-.049	-.051	-.039	.176	.129	.059	.128	33.30	.414
K	.162	.117	.015	-.193	-.112	-.059	.031	.137	-.013	.072	41.60	.431
C	.138	.075	.008	-.332	-.093	.024	.013	.249	-.036	.046	45.10	.553
Ap	.113	.025	.000	-.299	-.111	.010	.016	.284	-.101	.087	48.30	.557
An	.047	-.008	-.044	-.204	-.113	.026	-.028	.242	.003	.092	48.90	.437
Sy	.114	-.044	.009	-.134	-.136	.098	.023	.214	-.010	.109	37.40	.423
Ev	.009	.097	-.096	-.156	-.199	-.019	.200	.066	.027	.091	48.50	.430
K	.234	.119	.095	-.309	-.076	-.043	-.002	.222	-.056	.209	34.70	.588
C	.025	.058	.040	-.416	-.095	.058	.026	.254	.027	.204	40.30	.644
Ap	.030	.020	-.007	-.377	.012	.063	.022	.246	.023	.209	37.40	.581
An	.028	.001	-.027	-.396	-.068	-.008	.082	.251	.064	.134	46.50	.589
Sy	-.037	-.001	-.030	-.257	-.226	-.008	.148	.146	.012	.096	57.50	.482
Ev	.128	.006	-.028	-.058	-.126	-.043	.153	.015	.010	.017	45.90	.280
K	.078	.148	-.031	-.050	-.101	-.101	.181	.053	.027	.050	36.80	.360
C	.092	.104	-.003	-.346	-.093	-.023	.128	.239	-.037	.102	41.40	.578
Ap	.145	.096	.005	-.334	-.067	.031	.122	.222	-.047	.154	33.00	.581
An	.049	.170	-.029	-.244	-.062	.127	.070	.225	.019	.146	25.70	.517
Sy	.007	.113	-.006	-.106	-.123	-.005	.132	.111	.047	.146	33.70	.370
Ev	-.015	.018	-.054	-.144	-.153	-.006	.169	.100	-.018	.139	47.50	.385

TABLE 60

Summary of Significant Beta Weights

Grade 9

	Verbal Reasoning	Number Facility	Flexibility of Closure	Fluency	Speed of Closure	Memory	Figural Adaptive Flexibility	Figural Reasoning	Syllogistic Reasoning	Uninterpreted
Knowledge:										
AS =	$-B_1$	$-B_2$				B_6				
G =	$-B_1$		$-B_3$			B_6			B_9	$-B_{10}$
L =	$-B_1$			$-B_4$		B_6			B_9	
HG =	$-B_1$	$-B_2$		$-B_4$		B_6				$-B_{10}$
Comprehension:										
AS =	$-B_1$		$-B_3$	$-B_4$						
G =	[illegible]		$-B_3$	$-B_4$						
L =	$-B_1$	$-B_2$		$-B_4$		B_6	B_7			
HG =	$-B_1$			$-B_4$	$-B_5$	B_6	B_7			
Application:										
AS =	$-B_1$		$-B_3$	$-B_4$						
G =	$-B_1$		$-B_3$						$-B_9$	
L =	$-B_1$	$-B_2$		$-B_4$			B_7			
HG =	$-B_1$			$-B_4$			B_7		B_9	
Analysis:										
AS =	$-B_1$		$-B_3$							
G =	$-B_1$		$-B_3$					B_8		
L =	$-B_1$	$-B_2$		$-B_4$						
HG =	$-B_1$			$-B_4$		B_6				
Synthesis:										
AS =	$-B_1$	$-B_2$							B_9	
G =	$-B_1$			$-B_4$	$-B_5$	$-B_6$				B_{10}
L =	$-B_1$	$-B_2$		$-B_4$		B_6				$-B_{10}$
HG =	$-B_1$					B_6				
Evaluation:										
AS =	$-B_1$	$-B_2$			B_5					B_{10}
G =	$-B_1$	$-B_2$				$-B_6$				$-B_{10}$
L =	$-B_1$	$-B_2$		$-B_4$		B_6			B_9	
HG =	$-B_1$	$-B_2$	$-B_3$							

TABLE 60a

Summary of Predictors of Level Scores, over Forms
by Grades with Significant Beta Weights

Grade 9

	Reasoning	Verbal Reasoning	Number Facility	Flexibility of Closure	Fluency	Memory	Speed of Closure	Syllogistic Reasoning	Figural Adaptive Flexibility	Figural Reasoning	Uninterpreted
Knowledge:											
(AS)		-.43V	-.11N			+.18M					
(G)		-.43V		-.12Cf		+.18M		+.12Rs			+.13U
(L)		-.43V			-.10F	+.24M		+.10Rs			
(EG)		-.39V	-.14N		-.11F	+.23M					-.12U
Comprehension:											
(AS)		-.35V		-.10Cf	-.19F						
(G)		-.43V		-.25Cf	-.11F						
(L)		-.47V	-.17N		-.19F	+.12M			+.10Xa		
(EG)		-.51V			-.20F	+.15M	-.10Cs		+.10Xa		
Application:											
(AS)		-.42V		-.10Cf	-.16F						
(G)		-.55V		-.12Cf				-.11Rs			
(L)		-.48V	-.10N		-.16F				+.10Xa		
(EG)		-.46V			-.19F			+.09Rs	+.21Xa		
Analysis:											
(AS)		-.28V		-.13Cf							
(G)		-.41V		-.21Cf						+.14Fr	
(L)		-.41V	-.12N		-.14F						
(EG)		-.35V			-.15F	+.19M					
Synthesis:											
(AS)		-.16V	-.11N					+.12Rs			
(G)		-.22V			-.16F	-.12M	-.20Cs				+.12U
(L)		-.39V	-.19N		-.17F	+.08M					-.16U
(EG)		-.36V				+.18M					
Evaluation:											
(AS)		-.29V	-.16N								+.12U
(G)		-.38V	-.18N			+.15M	+.10Cs				
(L)		-.26V	-.19N		-.19F	+.11M		+.12Rs			
(EG)		-.42V	-.10N		-.14F						

TABLE 61

Summary of Significant Beta Weights

Grade 10

	Reasoning	Number Facility	Fluency	Verbal Reasoning	Flexibility of Closure	Semantic Redefinition	Deductive Reasoning	Memory	Fluency in Making Verbal Transformations	Speed of Closure
Knowledge:										
AS =	$-B_1$	B_2		$-B_4$			B_7			
G =	$-B_1$	B_2		$-B_4$			B_7	B_8		
L =	$-B_1$	B_2		$-B_4$			B_7	B_8		B_{10}
EG =	$-B_1$	B_2		$-B_4$			B_7	B_8	B_9	B_{10}
Comprehension:										
AS =	$-B_1$	B_2		$-B_4$			B_7			
G =	$-B_1$	B_2		$-B_4$	B_5		B_7			
L =	$-B_1$	B_2		$-B_4$			B_7			
EG =	$-B_1$	B_2		$-B_4$			B_7			
Application:										
AS =	$-B_1$	B_2		$-B_4$			B_7			
G =	$-B_1$	B_2		$-B_4$	B_5	$-B_6$	B_7		$-B_9$	
L =	$-B_1$	B_2		$-B_4$			B_7	B_8		
EG =	$-B_1$	B_2		$-B_4$			B_7			
Analysis:										
AS =	$-B_1$	B_2		$-B_4$			B_7			
G =	$-B_1$			$-B_4$			B_7			
L =	$-B_1$	B_2		$-B_4$			B_7			
EG =	$-B_1$	B_2		$-B_4$			B_7			
Synthesis:										
AS =		B_2		$-B_4$			B_7			
G =	$-B_1$		B_3	$-B_4$	B_5	$-B_6$	B_7	$-B_8$		
L =	$-B_1$	B_2	B_3	$-B_4$			B_7			
EG =		B_2		$-B_4$			B_7			
Evaluation:										
AS =	$-B_1$	B_2		$-B_4$			B_7	B_8		
G =	$-B_1$	B_2		$-B_4$			B_7			B_{10}
L =		B_2	B_3	$-B_4$			B_7		B_9	
EG =	$-B_1$	B_2		$-B_4$	B_5		B_7			

TABLE 61a

Summary of Predictors of Level Scores, over Forms
by Grades with Significant Beta Weights

Grade 10

	Reasoning	Verbal Reasoning	Number Facility	Flexibility of Closure	Fluency	Memory	Speed of Closure	Fluency in Making Verbal Transformations	Deductive Reasoning	Semantic Redefinition
Knowledge:										
(AS)	-.22R	-.36V	+.26N						+.24Rd	
(G)	-.21R	-.34V	+.23N			+.09M	+.16Cs		+.34Rd	
(L)	-.21R	-.32V	+.22N			+.13M			+.36Rd	
(EG)	-.14R	-.29V	+.25N			+.23M	+.10Cs	+.12FVT	+.23Rd	
Comprehension:										
(AS)	-.19R	-.15V	+.27N						+.23Rd	
(G)	-.27R	-.30V	+.24N	+.10Cf					+.31Rd	
(L)	-.24R	-.34V	+.29N						+.35Rd	
(EG)	-.21R	-.34V	+.13N						+.28Rd	
Application:										
(AS)	-.25R	-.27V	+.17N						+.23Rd	
(G)	-.25R	-.35V	+.09N	+.16Cf				-.12FVT	+.43Rd	-.11Re
(L)	-.24R	-.36V	+.23N			+.14M			+.36Rd	
(EG)	-.21R	-.31V	+.18N						+.28Rd	
Analysis:										
(AS)	-.15R	-.19V	+.14N						+.15Rd	
(G)	-.20R	-.29V							+.19Rd	
(L)	-.16R	-.35V	+.19N						+.26Rd	
(EG)	-.19R	-.17V	+.11N						+.29Rd	
Synthesis										
(AS)		-.24V	+.16N						+.13Rd	
(G)	-.16R	-.24V			+.12F	-.13M			+.15Rd	-.12Re
(L)	-.18R	-.34V	+.17N		+.10F				+.18Rd	
(EG)		-.24V	+.14N						+.20Rd	
Evaluation										
(AS)	-.17R	-.18V	+.19N			+.12M	+.11Cs		+.16Rd	
(G)	-.14R	-.32V	+.22N					+.16FVT	+.18Rd	
(L)		-.26V	+.22N		+.13F				+.13Rd	
(EG)	-.21R	-.21V	+.11N	+.11Cf					+.19Rd	

TABLE 62

Summary of Significant Beta Weights

Grade 11

	Reasoning	Number Facility	Deductive Reasoning	Speed of Closure	Verbal Reasoning	Fluency	Memory	Flexibility of Closure	Fluency in Making Verbal Transformations	Sensitivity to Problems
Knowledge:										
AS =	B_1				B_5	B_6	B_7			
G =	B_1	B_2			B_5		B_7			
L =	B_1	B_2			B_5		B_7			
EG =	B_1	B_2			B_5	B_6	B_7			
Comprehension:										
AS =	$-B_1$	B_2		$-B_4$	B_5	B_6	B_7		B_9	
G =	$-B_1$	B_2	B_3		B_5			$-B_8$		
L =	$-B_1$	B_2			B_5					
EG =	$-B_1$	B_2			B_5	B_6	B_7			
Application:										
AS =	$-B_1$	B_2			B_5	B_6	B_7		B_9	
G =	$-B_1$	B_2	B_3		B_5			$-B_8$	B_9	
L =	$-B_1$	B_2			B_5			$-B_8$	B_9	
EG =	$-B_1$	B_2			B_5	B_6	B_7	$-B_8$		
Analysis:										
AS =	$-B_1$	B_2			B_5	B_6		$-B_8$	B_9	
G =	$-B_1$	B_2			B_5			$-B_8$		
L =	$-B_1$	B_2			B_5					
EG =	$-B_1$	B_2			B_5	B_6				
Synthesis:										
AS =	$-B_1$	B_2			B_5			$-B_8$	B_9	
G =	$-B_1$				B_5		B_7			
L =	$-B_1$	B_2			B_5	B_6	B_7			
EG =	$-B_1$	B_2			B_5					
Evaluation										
AS =	$-B_1$				B_5	B_6	B_7			
G =	$-B_1$	B_2	B_3		B_5	B_6	B_7			
L =	$-B_1$	B_2			B_5					
EG =	$-B_1$	B_2	B_3	B_4	B_5		B_7			

TABLE 62a

Summary of Predictors of Level Scores, over Forms
by Grades with Significant Beta Weights

Grade 11

	Reasoning	Verbal Reasoning	Number Facility	Flexibility of Closure	Fluency	Memory	Speed of Closure	Fluency in Making Verbal Transformations	Deductive Reasoning	Sensitivity to Problems
Knowledge:										
(AS)	-.23R	+.34V			+.10F	+.25M				
(G)	-.31R	+.19V	+.17N			+.22M				
(L)	-.31R	+.24V	+.14N			+.12M				
(EG)	-.23R	+.22V	+.21N		+.14F	+.23M				
Comprehension:										
(AS)	-.22R	+.16V	+.10N		+.11F	+.12M	-.10Cs	+.11FVT		
(G)	-.34R	+.25V	+.15N	-.20Cf					+.11Rd	
(L)	-.32R	+.28V	+.16N							
(EG)	-.31R	+.32V	+.15N		+.14F	+.12M				
Application:										
(AS)	-.25R	+.22V	+.12N		+.10F	+.16M		+.14FVT		
(G)	-.38R	+.23V	+.14N	-.12Cf				+.19FVT	+.11Rd	
(L)	-.34R	+.30V	+.13N	-.11Cf				+.10FVT		
(EG)	-.32R	+.28V	+.18N	-.10Cf	+.11F	+.15M				
Analysis:										
(AS)	-.20R	+.14V	+.10N	-.10Cf	+.13F			+.13FVT		
(G)	-.28R	+.18V	+.12N	-.21Cf						
(L)	-.31R	+.27V	+.11N							
(EG)	-.24R	+.17V	+.14N		+.11F					
Synthesis:										
(AS)	-.13R	+.19V	+.13N							
(G)	-.20R	+.19V		-.23Cf		-.11M		+.12FVT		
(L)	-.22R	+.34V	+.14N		+.14F	+.10M				
(EG)	-.17R	+.23V	+.15N							
Evaluation:										
(AS)	-.18R	+.12V			+.17F	+.13M				
(G)	-.17R	+.15V	+.13N		+.17F	+.17M			+.14Rd	
(L)	-.21R	+.13V	+.15N							
(EG)	-.19R	+.22V	+.17N			+.12M	+.10Cs		+.14Rd	

TABLE 63

Summary of Significant Beta Weights

Grade 12

	Flexibility of Closure	Number Facility	Speed of Closure	Verbal Reasoning	Fluency	Memory	Uninterpreted	Reasoning	Fluency in Making Verbal Transformations	Syllogistic Reasoning
Knowledge:										
AS =	B_1			$-B_4$			B_7	B_8		
G =	B_1	B_2		$-B_4$	$-B_5$			B_8		B_{10}
L =	B_1	B_2	B_3	$-B_4$				B_8		
EG =		B_2					B_7			B_{10}
Comprehension:										
AS =	B_1			$-B_4$				B_8		
G =	B_1			$-B_4$	$-B_5$			B_8		
L =				$-B_4$	$-B_5$			B_8		B_{10}
EG =		B_2		$-B_4$			B_7	B_8		B_{10}
Application:										
AS =	B_1	B_2		$-B_4$				B_8		
G =	B_1			$-B_4$	$-B_5$			B_8	$-B_9$	B_{10}
L =				$-B_4$				B_8		B_{10}
EG =	B_1	B_2		$-B_4$			B_7	B_8		B_{10}
Analysis:										
AS =	B_1			$-B_4$				B_8		
G =				$-B_4$	$-B_5$			B_8		
L =				$-B_4$				B_8		B_{10}
EG =		B_2		$-B_4$		B_6		B_8		B_{10}
Synthesis:										
AS =										
G =	B_1			$-B_4$	$-B_5$			B_8		
L =				$-B_4$	$-B_5$		B_7	B_8		B_{10}
EG =					$-B_5$		B_7	B_8		
Evaluation:										
AS =		B_2	$-B_3$		$-B_5$		B_7	B_8		B_{10}
G =				$-B_4$	$-B_5$		B_7			
L =	B_1				$-B_5$		B_7			
EG =				$-B_4$	$-B_5$		B_7			B_{10}

TABLE 63a

Summary of Predictors of Level Scores, over Forms
by Grades with Significant Beta Weights

Grade 12

	Reasoning	Verbal Reasoning	Number Facility	Flexibility of Closure	Fluency	Memory	Speed of Closure	Fluency in Making Verbal Transformations	Syllogistic Reasoning	Uninterpreted
Knowledge:										
(AS)	+.13R	-.11V		+.22Cf					+.13Rs	+.15U
(G)	+.14R	-.19V	+.12N	+.16Cf	-.11F					
(L)	+.22R	-.31V	+.12N	+.13Cf			+.10Cs		+.21Rs	
(EG)			+.15N							+.18U
Comprehension:										
(AS)	+.15R	-.18V		+.16Cf						
(G)	+.25R	-.33V		+.14Cf	-.10F					
(L)	+.25R	-.42V			-.10F				+.20Rs	
(EG)	+.24R	-.35V	+.10N						+.10Rs	+.13U
Application:										
(AS)	+.19R	-.13V	+.13N	+.24Cf					+.14Rs	
(G)	+.28R	-.30V		+.11Cf	-.11F			-.10FVT		
(L)	+.25R	-.38V							+.21Rs	
(EG)	+.22R	-.33V	+.10N	+.14Cf					+.15Rs	+.12U
Analysis:										
(AS)	+.11R	-.11V		+.20Cf						
(G)	+.24R	-.20V			-.11F					
(L)	+.25R	-.40V								
(EG)	+.22R	-.24V	+.17N			+.13M			+.15Rs	
Synthesis:										
(AS)	+.23R								+.17Rs	
(G)	+.21R	-.13V		+.11Cf	-.14F				+.11Rs	
(L)	+.15R	-.26V			-.23F					+.15U
(EG)	+.11R		+.11N		-.12F				+.15Rs	+.13U
Evaluation:										
(AS)	+.13R		+.14N				-.18Cs		+.13Rs	+.18U
(G)		-.16V			-.20F					+.20U
(L)				+.13Cf	-.13F					+.15U
(EG)		-.14V			-.15F				+.14Rs	+.17U

TABLE 64

Named KIT Factors by Grade

	Grade 9	Grade 10	Grade 11	Grade 12
I.	Verbal Reasoning (V)	Reasoning (R)	Reasoning (R)	Flexibility of Closure (Cf)
II.	Number Facility (N)	Number facility (N)	Number facility (N)	Number Facility (N)
III.	Flexibility of Closure (Cf)	Fluency (F)	Deductive Reasoning (Rd)	Speed of Closure (Cs)
IV.	Fluency (F)	Verbal Reasoning (V)	Speed of Closure (Cs)	Verbal Reasoning (V)
V.	Speed of Closure (Cs)	Flexibility of Closure (Cf)	Verbal Reasoning (V)	Fluency (F)
VI.	Memory (M)	Semantic Redefinition (Re)	Fluency (F)	Memory (M)
VII.	Figural Adaptive Flexibility (Xa)	Deductive Reasoning (Rd)	Memory (M)	Uninterpreted (U)
VIII.	Figural Reasoning (FR)	Memory (M)	Flexibility of Closure (Cf)	Reasoning (R)
IX.	Syllogistic Reasoning (Rs)	Fluency in Making Verbal Transformation (FVT)	Fluency in Making Verbal Transformation (FVT)	Fluency in Making Verbal Transformations (FVT)
X.	Uninterpreted (U)	Speed of Closure (Cs)	Sensitivity to Problems (Sep)	Syllogistic Reasoning (Rs)

Comparisons by grade and form. For Grade 9 - Atomic Structure, Verbal Reasoning is the only ability common to all process levels, although a fluency factor, either Flexibility of Closure, Number Facility, Fluency, or Speed of Closure appears at each process level. The combination of Verbal Reasoning, Number Facility and Memory factors in the equation for Knowledge is probably the most logical one in terms of the process.

For Grade 9 - Glaciers, Verbal Reasoning and some kind of fluency again appear at each level for this test form. Although the appearance of the Memory factor in the Knowledge equation is logical, its appearance in the Synthesis and Evaluation process equations is difficult to explain on the basis of the taxonomy structure. However, the reader will recall that the simplex ordering of subtests according to empirically determined complexity was Knowledge, Synthesis, Evaluation, Comprehension, etc. Consequently, the appearance of the Memory factor in the Knowledge, Synthesis, and Evaluation equations is not unreasonable.

For Grade 9 - Lisbon Earthquake, the equations show more consistency than do the others for grade nine. Verbal Reasoning and Fluency appear in all equations; Number Facility appears in all except Knowledge; and Memory appears in four of the six equations. At the ninth-grade level, this form was the only one for which the theoretical order and empirically determined complexity order were identical.

For Grade 9 - Stages of Economic Growth, Verbal Reasoning appears in all equations and Fluency appears in the equations for the first four levels. The Synthesis equation contains only Verbal Reasoning and Memory which do not agree particularly with the definition of the process. However, Synthesis was ordered by the simplex immediately after Knowledge.

For Grade 10 - Atomic Structure, the equations for Knowledge, Comprehension, Application and Analysis contain the same four predictors, Reasoning, Number Facility, Verbal Reasoning, and Deductive Reasoning. The more general Reasoning factor drops out of the Synthesis equation. Speed of Closure and Memory appear in the Evaluation equation in addition to the four factors present at the lower levels.

For Grade 10 - Glaciers, Reasoning, Verbal Reasoning and Deductive Reasoning are common to all process levels. Fluency factors appear at several levels, in several combinations.

For Grade 10 - Lisbon Earthquake, four factors--three reasoning factors and Number Facility--appear in the equations for Knowledge through Synthesis. The equation for Evaluation lacks the general reasoning factor and has an additional fluency factor. Memory appears in the equations for Knowledge and Application.

For Grade 10 - Stages of Economic Growth, the same four factors--Reasoning, Number Facility, Verbal Reasoning, and Deductive Reasoning--appear in five of six equations; the general reasoning factor is the only one of the four missing from the sixth equation. Memory and two fluency factors appear in addition to the four at the Knowledge level.

For Grade 11 - Atomic Structure, Reasoning and Verbal Reasoning appear in all equations; Fluency appears in all equations except that for Synthesis and Number Facility is present in all except that for Knowledge and Evaluation. Memory appears in all equations except Analysis and Synthesis.

For Grade 11 - Glaciers, Reasoning and Verbal Reasoning appear in all equations. Number Facility appears at all levels except Synthesis; Flexibility of Closure enters at the Comprehension level and remains in all subsequent equations. Fluency in Making Verbal Transformations appears in the equations for Application and Synthesis.

For Grade 11 - Lisbon Earthquake, three factors--Reasoning Number Facility and Verbal Reasoning--appear in all equations. Memory also appears in the Knowledge and Synthesis equations, and the Application equation includes two additional fluency factors.

For Grade 11 - Stages of Economic Growth, the set of equations appears to contain a progression of factors over process levels. The same five factors define Knowledge and Comprehension; these five plus a fluency factor define Application; the number of factors drops to four for Analysis, Memory and Flexibility of Closure dropping out; and for Synthesis, Reasoning, Verbal Reasoning and Number Facility are the only factors in the equation. The equation for Evaluation conflicts with the trend because it contains six factors, two of which do not appear in any of the other equations.

For Grade 12 - Atomic Structure, fluency and reasoning factors define all process levels except for Synthesis where Reasoning is the only significant factor.

For Grade 12 - Glaciers, fluency and reasoning factors appear in all six equations. All equations except Evaluation include two reasoning factors and all equations except those for Analysis and Evaluation include two fluency factors.

For Grade 12 - Lisbon Earthquake, the equations show some progression. Knowledge is defined by both fluency and reasoning factors; fluency becomes less important for Comprehension and for Application and Analysis, only reasoning factors appear. There are incongruities in the equations for Synthesis which includes Fluency, and especially for Evaluation which depends only on Flexibility of Closure.

For Grade 12 - Stages of Economic Growth, the equations for Knowledge through Analysis contain four common factors, three reasoning factors and one fluency factor. The presence of the unidentified factor in five of the six equations complicates the interpretation of these equations.

Comparisons by grade over forms. For these comparisons, a cognitive factor-process level relationship is inferred only if two or more of the same process level scores are significantly predicted by the same cognitive factor.

Grade Nine. The following equations can be written for the process scores:

	Common Factors							Special Factors
	R	V	N	Cf	F	M	Cs	
Knowledge		V +	N +		F +	M +		Rs + U
Comprehension		V +		Cf +	F +	M +		Xa
Application		V +		Cf +	F +	M +		Rs
Analysis		V +		Cf +	F			
Synthesis		V +		Cf +	F +	M +		U
Evaluation		V +	N +			M +		U

These are certainly not formal regression equations. They should be interpreted as in the following example: The abilities important for the Comprehension process are Verbal Reasoning, Flexibility of Closure, Fluency, Memory, and Figural Adaptive Flexibility.

Verbal Reasoning is the only ability common to all processes at grade nine. Fluency plays a significant role for all levels except Evaluation. Other abilities which are significantly related to performance are Flexibility of Closure, Number Facility, and the Unidentified Factor. In general, the equations for grade nine do not present any clear patterns.

Grade Ten. The following equations were determined for grade ten:

	Common Factors							Special Factors
	R	V	N	Cf	F	M	Cs	
Knowledge	R +	V +	N +			M +	Cs +	Rd
Comprehension	R +	V +	N +					Rd
Application	R +	V +	N +					Rd
Analysis	R +	V +	N +					Rd
Synthesis	R +	V +	N +		F +			Rd
Evaluation	R +	V +	N +					Rd

The equations for grade ten are much more homogeneous than those for grade nine. Four factors--Reasoning, Number Facility, Verbal Reasoning, and Deductive Reasoning--are significantly related to all process levels. In addition, Memory and Speed of Closure appear only at the Knowledge level, where they could logically be expected and thereby lending support to theoretical speculations about the nature of that level. Fluency, the process of putting various aspects of knowledge together in ways to form new products, appears as a significant predictor for Synthesis.

Grade Eleven. The following equations were determined for grade eleven:

	Common Factors							Special Factors
	R	V	N	Cf	F	M	Cs	
Knowledge	R +	V +	N +		F +	M		
Comprehension	R +	V +	N +		F +	M		
Application	R +	V +	N +	Cf +	F +			FVT
Analysis	R +	V +	N +	Cf +	F			
Synthesis	R +	V +	N +			M		
Evaluation	R +	V +	N +		F +	M +		Rd

The patterns for grade eleven, while more homogeneous than those for grade nine, are not as clear as those for grade ten. Three factors, Reasoning, Number Facility and

Verbal Reasoning appear in the equations for all levels; Memory and Fluency appear in five of six equations but Memory drops out on Analysis and Fluency drops out on Synthesis. Flexibility of Closure is significant for Application and Analysis; in addition, Fluency in Making Verbal Transformations is significantly related to Application. In Guilford's classification, Flexibility of Closure is said to represent convergent production of figural transformations. The factor FVT would probably involve divergent production of verbal transformations. Hence, these relationships appear to be logical ones because Application and Analysis involve transformations but Application depends on divergent production which Analysis does not. The appearance of Deductive Reasoning for Evaluation is logically effected.

Grade Twelve. The following equations were determined for grade twelve.

	Common Factors							Special Factors
	R	V	N	Cf	F	M	Cs	
Knowledge	R +	V +	N +	Cf +				Rs + U
Comprehension	R +	V +		Cf +	F +			Rs
Application	R +	V +	N +	Cf +				Rs
Analysis	R +	V +						Rs
Synthesis	R +	V +			F +			U
Evaluation		V +			F +			Rs + U

Again, the differential aspects of the equations should be noted. Except for the Verbal Reasoning factor, none appears in all six equations. Reasoning and Syllogistic Reasoning, appear at five levels; Reasoning drops out for Evaluation and Syllogistic Reasoning drops out for Synthesis.

Comparisons by form over grades. The following predictor equations were derived by identifying those factors which significantly predicted the same process level at two or more grade levels on each form. The reader is reminded that the same factor designation at different grade levels does not imply the same factor components.

Atomic Structure. The following equation can be written for the Atomic Structure form:

	Common Factors							Special Factors
	R	V	N	Cf	F	M	Cs	
Knowledge	R +	V +	N +			M		
Comprehension	R +	V +	N +	Cf +	F			
Application	R +	V +	N +	Cf +	F			
Analysis	R +	V +	N +	Cf				
Synthesis	R +	V +	N					
Evaluation	R +	V +	N +			M +	Cs	

The Verbal Reasoning, general Reasoning, and Number Facility factors are common to all process levels. The presence of additional fluency and/or reasoning factors for the lower levels is a logical expectation. Memory, appearing in the equation for Knowledge is to be expected; the presence of Memory and Speed of Closure in the Evaluation equation could, perhaps, be attributed to the use of external criteria in the evaluative process.

Glaciers. The following equation can be written for the Glaciers form:

	Common Factors							Special Factors
	R	V	N	Cf	F	M	Cs	
Knowledge	R +	V +	N +	Cf +		M		
Comprehension	R +	V +	N +	Cf +	F +			Rd
Application	R +	V +	N +	Cf +				Rd + FVT
Analysis	R +	V +		Cf				
Synthesis	R +	V +		Cf +	F +	M		
Evaluation	R +	V +	N +		F +			Rd

The presence of six factors in the equations for four of the levels is noteworthy. Reasoning and fluency factors are about equally important at all levels.

Lisbon Earthquake. The following equations can be written for the Lisbon Earthquake form:

	Common Factors							Special Factors
	R	V	N	Cf	F	M	Cs	
Knowledge	R +	V +	N			M +		Rs
Comprehension	R +	V +	N +		F			
Application	R +	V +	N					
Analysis	R +	V +	N					
Synthesis	R +	V +	N +			M		
Evaluation		V +	N +		F			

The levels from Knowledge through Analysis have four similar factors with Memory added to the Knowledge equation and Fluency added to the Comprehension equation. Evaluation appears to be predominantly a fluency factor rather than a reasoning factor.

<u>Stages of Economic Growth</u>. The following equations can be written for the Stages of Economic Growth form:

	Common Factors							Special Factors
	R	V	N	Cf	F	M	Cs	
Knowledge	R +	V +	N +		F +	M		
Comprehension	R +	V +	N +		F			
Application	R +	V +	N +	Cf +				Rd
Analysis	R +	V +	N +			M		
Synthesis		V +	N					
Evaluation	R +	V +	N +	Cf +				Rd

Five equations contain the same four factors--Verbal Reasoning, Number Facility, Reasoning and Deductive Reasoning. Additional fluency factors appear in the equations for Knowledge, Comprehension, Application and Evaluation; and Memory appears in the Knowledge and Analysis equations. The Synthesis equation contains only two factors, Verbal Reasoning and Number Facility.

<u>Comparisons over grades and forms</u>. The following discussion deals with similarities and differences between equations for each process level for each grade considered over forms. To be included in the equation for a grade, the cognitive factor had to appear as a significant predictor of the process for two or more forms. The first equation in each set is for grade nine, the second for grade ten, etc.

<u>Knowledge</u>

K_9	=		V +	N +		F +	M +		Rs +	U
K_{10}	=	R +	V +	N +				Cs +	Rd	
K_{11}	=	R +	V +	N +		F +	M			
K_{12}	=	R +	V +	N +	Cf +				Rs +	U

Verbal Reasoning and Number Facility are significant predictors of Knowledge at all grades. Memory, Syllogistic Reasoning and General Reasoning are important at three grade levels, memory drops out for grade twelve, deductive reasoning for grade eleven and General Reasoning for grade nine. Fluency appears at two grade levels, as do the unidentified factors for grades nine and twelve.

Comprehension

$$
\begin{array}{llllllllll}
C_9 & = & & V\ + & & Cf\ + & F\ + & M\ + & & Xa \\
C_{10} & = & R\ + & V\ + & N\ + & & & & & Rd \\
C_{11} & = & R\ + & V\ + & N\ + & & & M\ + & & Rd \\
C_{12} & = & R\ + & V\ + & & Cf\ + & F\ + & & & Rs
\end{array}
$$

Comprehension involves the ability to understand both literal and nonliteral statements, interpretations and/or extrapolation. The appearance of the reasoning factors at all grade levels and the Flexibility of Closure and Fluency factors seem to agree logically with the definition of comprehension, even though the same factors do not all appear at each grade level.

Application

$$
\begin{array}{lllllllll}
Ap_9 & = & & V\ + & & Cf\ + & F\ + & & Rs + Xa \\
Ap_{10} & = & R\ + & V\ + & N\ + & & & & Rd \\
Ap_{11} & = & R\ + & V\ + & N\ + & Cf\ + & F\ + & & FVT \\
Ap_{12} & = & R\ + & V\ + & N\ + & Cf\ + & & & Rs
\end{array}
$$

The pattern of prediction for Application is slightly more clear than that for Comprehension. Verbal Reasoning, General Reasoning and Syllogistic Reasoning, Flexibility of Closure and Fluency factors appear as the most important predictors of Application. Again, these factors generally agree with the definition of the behavior being predicted.

Analysis

$$
\begin{array}{lllllllll}
An_9 & = & & V\ + & & Cf\ + & F & & \\
An_{10} & = & R\ + & V\ + & N\ + & & & & Rd \\
An_{11} & = & R\ + & V\ + & N\ + & Cf\ + & F & & \\
An_{12} & = & R\ + & V\ + & & & & & Rs
\end{array}
$$

Analysis is defined to be an ability wherein a communication is separated into its constituent parts such that the communication is made more clear. Thus, the abilities measured by the several reasoning tests should be predictors of this behavior. This was primarily the case for grades ten, eleven, and twelve. As Analysis is one of the higher levels, the possibility exists that the abilities which define the process are not well developed for students in grade nine.

Synthesis

S_9	=	V + N +	F + M +	U
S_{10}	=	R + V + N +	F +	Rd
S_{11}	=	R + V + N +	M	
S_{12}	=	R + V +	F +	U

Synthesis, defined as putting together of elements to form a new communication involves Verbal Reasoning and other kinds of reasoning, and Fluency. The significance of Number Facility and Memory in predicting Synthesis is difficult to explain. The possibility exists that it is the cumulative effect of these abilities appearing in a higher level process.

Evaluation

E_9	=	V + N +	M +	U
E_{10}	=	R + V + N +		Rd
E_{11}	=	R + V + N +	F + M +	Rd
E_{12}	=	V +	F	U + Rs

Predominant in the prediction of Evaluation is the number of dissimilarities which appear over grade levels. Consistency with respect to Verbal Reasoning and Syllogistic and Deductive Reasoning provides a logical identification with the process, but is clouded by the ability factors which appear at only one or two grade levels.

Trends over levels, forms, and grades. The following sections present a summary and interpretation of changes in beta weights which occur in the regression equations. The method of detecting trends was applied only to the six factors which were common for all four grades and the two which were common for grades ten through twelve. For each factor separately, the average beta weights were computed for each process level within grade and over forms. These averages were then averaged over levels and also over grades. These averages of averages should reveal trends in magnitude of beta weights over grades and levels. The data on which these are based appear in Tables 60a, 61a, 62a, and 63a.

The Reasoning factor is common to the three upper grades. The beta weights are substantial over all grades and levels. The average over levels was highest for grade eleven. The averages over grades increase from Knowledge through Application and decrease from Analysis through Evaluation.

The beta weights for Verbal Reasoning are substantial over all grades and levels. The averages over levels decrease

in size from grades nine to twelve. There is a tendency for the average weights over grades to decrease from Knowledge through Evaluation.

For Number Facility the only substantial weights appear at grade ten for Knowledge and Comprehension. The average weights over levels are greatest for grades ten and eleven. The weights, over grades, are generally small, but there is a decrease in average weight from Knowledge through Synthesis, with a slight increase in the average weight for Evaluation.

Flexibility of Closure has no substantial beta weights over grades or levels. The averages over levels are approximately the same size for grades nine, eleven and twelve, but there is a noticeable drop in the weight for grade ten. The average over grades for Knowledge is low but the weights increase and remain constant for Comprehension, Application and Analysis and drop to a negligible amount for Synthesis and Evaluation.

The loadings for the Fluency factor are quite small, The average weight over levels for grade ten was nearly zero; those for grades nine, eleven, and twelve were small and approximately equal. The average weights for levels over grades were small and showed no pattern.

The beta weights for the Memory factor were not substantial. The averages over levels for grades nine through twelve showed a generally decreasing pattern, implying that Memory was of increasingly less importance as grade level increased. The averages over grades for levels revealed a predictable pattern; the average was largest for Knowledge, dropped to a negligible amount for Comprehension through Synthesis and increased for Evaluation. One would expect Memory to be important in predicting Knowledge scores and, because the simplex analyzed ordered Evaluation low in the hierarchy, Evaluation scores too.

Speed of Closure had small beta weights over grades and levels. The averages over levels for grades nine and ten were the same and were larger than the comparable averages for grades eleven and twelve, which were of the same size. The averages over grades for levels were practically zero except for Knowledge and Evaluation.

The factor, Fluency in making Verbal Transformations, was common only to the three upper grades. Its beta weights were generally not substantial for levels or grades. The largest average over levels for grades appeared at grade eleven. The averages over grades for levels were small but constant for all levels.

Discussion

It was presumed that sixteen psychologically pure factor tests could be extracted from the thirty-seven cognitive aptitude tests from the KIT and that these would be common for all grade levels. But the factor analyses revealed that only six factors could be extracted which were common for all grades and that two others could be extracted which were common for grades ten through twelve. At each grade, ten factors were extracted, so at any grade level, as few as two and as many as four factors could be specific to the grade level. The composition of each of the eight common factors (six at four grade-levels and two at three grade-levels) was not similar from one grade-level to the next to the extent to which we had anticipated.

The failure to obtain the expected number of factors is probably attributable in part to the restriction of range of scores on some tests, because some were too difficult for the youngest age groups and too simple for the oldest age groups, and because of the possibility that the elemental factors are organized differently in students of different ages possibly because of maturational or experimental differences.

The consequence of failing to obtain the factors in the expected number and generality over grades seriously impeded the interpretation of the regressional analyses which were intended to reveal the similarity of process scores over content and grade and the changes in the factor structure of different process scores within grade, etc.

It was our expectation that the factor structure for a given level would be similar across contents and possibly over grades. It was our unstated expectation that (a) if all factors related to all level scores, then there would be systematic changes in the beta weights for each of the factors across the several levels, or (b) if all factors did not relate significantly to each level score, then as level of progress increased then new factors would appear in the regression equation. In short, the first hypothesis is that all factors appear at all levels but at each level the weighted composition of them differs and that the changes in weights for a given factor would be orderly through the levels. The second hypothesis is that as level increases the complexity does too and as a consequence additional factors will appear in the regression equations at a given level and will remain in the equations for all higher levels of process. Frankly, we did not know which hypothesis would more probably be supported by the data.

It is obvious that our failing to obtain the sixteen factors which we had anticipated and our obtaining the factors we did, drastically reduced the chances of determining which of the previously mentioned hypotheses about the factor structure of levels was more tenable. But even without regard to the factors which were obtained, the results of the simplex and circumplex analyses made it rather evident that the investigation of the psychological structure in terms of elemental aptitudes would be rather difficult methodologically and that whatever results were obtained would be equivocal. Specifically, the factor analyses which revealed the presence of content factors clouded the relationships between aptitude and process scores. The simplical analyses revealed for several forms and grades that Synthesis and Evaluation might more properly be regarded as coming between Knowledge and Comprehension in an ordering according to empirically determined complexity. This information was damaging on two counts. First, that suggested misordering was not consistent over forms and grades which implied that the several forms had different ability compositions for the same process and/or that the organization of abilities within individuals were associated with their age and grade placement. Second, there was a clear implication that some process levels, specifically Synthesis and Evaluation, were deficient inasmuch as the contradictions in the results of the simplex analyses could partially be explained on the basis of item deficiencies which were known or suspected prior to the analyses. The circumplex analyses suggested that the Knowledge and Evaluation items were regarded as the weakest of all that were written, there was some doubt that the regression analyses would reveal systematic changes in factor structure from level to level so long as a particular order was forced on the levels across all contents. In summary, the equivocal results of the regression analyses were not unexpected.

Nevertheless, the results do reveal some consistencies which were expected. Verbal Reasoning, Reasoning, and Number facility load substantially at nearly all levels. Memory loads most often and heavily on Knowledge, Synthesis and Evaluation. (This expectancy was generated on the basis of the simplex and circumplex analyses.) Speed of Closure had greater loadings on Knowledge and Evaluation; again in accordance with the expectancy resulting from testing the first two hypotheses.

D. Summary

This section dealt with testing hypotheses about the hierarchical structure of the taxonomy and the generality of

the processes; and determining the psychological structure of the six processes.

The hypothesis about the hierarchical structure was tested with two sets of data. Mean scores from subtests representing the processes were examined to determine if they decreased as the taxonomic level of the processes increased. The results were generally supportive; however, on the two forms which contained science content there were reversals at every grade level of the Synthesis and Evaluation means, and on one of these forms, the Comprehension and Application means were reversed at every grade level. This latter set of reversals is attributable to several Comprehension items which required information which was quite inconspicuously placed in the reading passage. The reversal of the Synthesis and Evaluation means is probably due to a combination of the high difficulty of the contents which made the principles difficult to grasp and of the Evaluation items which required the use of external criteria (not included in the reading passage) which upon later analysis seemed to be more a matter of commonsense than rigorous evaluative procedures.

The second set of evidence for testing the hierarchy was the intercorrelation matrices for subtest (process) scores for each form and grade level. These matrices were analyzed for simplex structure according to the Guttman Radex theory. The relationship between the taxonomy ordering according to complexity was then compared with the best empirically determined order according to a method devised by Kaiser. The theoretical order and the best empirical order were the same at all grade levels for the Lisbon Earthquake test. The Atomic Structure form contained only one discrepancy, Synthesis and Evaluation were placed between Knowledge and Comprehension and this occurred at only the ninth-grade level. The Stages of Economic Growth form contained only two discrepancies; in grade nine Synthesis, was placed between Knowledge and Comprehension, and at grade twelve, Evaluation was placed between Analysis and Synthesis. On the Atomic Structure form, Synthesis and Evaluation were empirically placed between Knowledge and Comprehension at every grade level.

The overall conclusion drawn from these two sets of evidence is that the hypothesis about the hierarchical structure of the taxonomy is neither completely confirmed nor disconfirmed. However, the tendency in the evidence is to support the hypothesis and this tendency is further strengthened because some of the data which are at variance with the hypothesis can be ascribed partially to faulty subtest items.

The second hypothesis dealt with the transcendence of processes over content. Two sets of data were used to investigate its validity. The six processes scores from each of the four forms for each grade level separately were intercorrelated and factor analyzed. The unrotated factor matrices were then inspected for the presence of process factors which appeared over the four forms. These analyses revealed the presence of a general factor and some mixed or process-content factors. Factors associated with each of Knowledge, Synthesis, and Evaluation gave the clearest evidence of the generality of process. However, the majority of factors were mixtures of both content and process which tends to discredit the hypothesis.

The second set of data consisted of intercorrelation matrices of the same subtest (process) scores for the four forms at each level separately. Each of the resulting twenty-four matrices were subjected to circumplex analysis. The requirement for a perfect circumplex is that column totals for the matrix be equal. None of the obtained matrices exhibited this property; however, many of them contained small differences among the column totals. Because no method was available with which to determine the significance of the relationship between the obtained values and the theoretically desirable values, subjective judgments were made about which of the matrices exhibited enough consistency to be regarded as quasi-circumplex. The most consistent matrices were those for Knowledge and Evaluation.

The evidence provided by the two sets of data was such that the hypothesis was regarded as being neither proved nor disproved. If a tendency exists in the data, then it is probably in the direction of confirmation. An adequate test of the hypothesis must await the development of methods for determining the significance of the relationship between an obtained matrix and one which exhibits circumplex structure, and until a more appropriate factor analysis technique is located or devised for treating presumptively hierarchical sets of data.

The final set of investigations dealt with the psychological structure of the taxonomy by identifying the factor composition of each of the subtests of the four taxonomy forms. The cognitive factor scores were obtained by administering thirty-seven tests from the Kit of Reference Tests for Cognitive Factors to students from grades nine through twelve. These were then intercorrelated and factor analyzed. Ten factors were extracted from each grade matrix. Six of them were regarded as being common to the upper three grades. But none of these eight factors was

identically determined at all grade levels. The remaining factors were common to two grades, specific to a grade, or were uninterpretable.

The nature and number of extracted factors were different from the expectation and seriously impeded the interpretation of the regression analyses which were subsequently calculated. The thirty-seven tests were chosen from the KIT on the assumption that they would yield sixteen psychologically pure factor tests. It was also assumed that each of these factors would have highly similar compositions over grade levels. The obtained factors discredited both assumptions.

Each student's score on each of the six subtests from each of the four taxonomy-type tests was predicted from his ten factor scores. The resulting regression equations were then analyzed to determine if the same factors appeared for each of the test forms at each grade level, whether there were systematic changes in the equations for process levels within and over grades, etc. No clear antecedent hypothesis had been formulated about the probable results. However, it was expected that one of the following statements would be descriptive of the results. All factors will load on each subtest (process) of all four test forms at all grade levels, and systematic changes in the beta weights will be noted from the Knowledge subtest through the Evaluation subtest and systematic changes in beta weights will be noted over grade levels. The competing expectation was that the regression equations would contain an increasing number of significant factors from Knowledge through Evaluation and that over grades the same order of appearance and continuance of factors in the regression equations would be noted. Neither one of these postulates was verified by the data. Some of the common factors had stable weights over grades and systematic changes over levels and vice versa, others held relatively constant values over levels and grades, and some changed over both simultaneously. Nevertheless, certain apparently consistent patterns did emerge. Due to the unsatisfactory number of common factors and their lack of identical composition over grade levels, the patterns which did appear are regarded as being undetermined but suggestive of the probable value of more rigorous subsequent analyses.

VII. SUMMARY AND RECOMMENDATIONS

This section contains a summary of the entire project and recommendations about where future research activities might be pointed to realize the greatest gains. These recommendations spring from problems which arose during the project and were not resolved.

A. Summary

Problem

The major purpose of the project was to determine the construct validity of the classification scheme presented in the Taxonomy of Educational Objectives, Handbook I: Cognitive Domain.

The classification system or taxonomy is intended for organizing educational objectives which describe cognitive behaviors. The taxonomy consists of six major levels; Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. Each of these major levels consists of sublevels.

The taxonomy was constructed to be hierarchical and cumulative. The basis of the hierarchy is the increasing complexity of the cognitive processes which appear in the taxonomy. The imputed order is from Knowledge, the simplest, to Evaluation, the most complex, in the sequence which appears above. The cumulative characteristic refers to the inclusiveness of each major level; specifically each major level contains all processes described at all levels which are placed lower than it in the structure. Therefore, the Analysis level is less complex than Synthesis and Evaluation; it is more complex than Knowledge, Comprehension, and Application; and it includes all processes that are implied by Knowledge, Comprehension, and Application, and it also includes novel processes with respect to those in the lower levels.

The phenomena which are classified by this system are the responses (behaviors) which test items (in the cognitive domain, as contrasted to the affective and psychomotor domains) are intended to evoke.

The implicit model of a response to a test item is, in its simplest form, that the response is the interaction of the application of a mental process to an unit of subject matter content.

The taxonomists state their belief that the cognitive processes, which are described in the taxonomy, transcend content; i.e., cognitive processes described by Comprehension constitute mental or cognitive operators which are equally applicable to many contents.

The Taxonomy was published in 1956 and since that time it has served to structure the conceptualization of educational problems and activities. It is used to classify educational objectives, to formulate the specifications for achievement tests, to develop sequential training programs which hopefully lead to the attainment of increasingly complex behavior patterns, etc. Despite these significant usages, few studies appear in the literature which deal pointedly with the validity of the taxonomic structure. This project was addressed to that topic. Three questions were considered.

1. Can empirical evidence be found to support or refute the imputed hierarchical structure?

2. Can empirical evidence be found to support or refute the imputed generality of the several cognitive processes?

3. Can each level of the structure be explained by more elemental cognitive aptitudes, and, if so, do the combinations or numbers of them change systematically from one major level to the next?

The question or hypothesis about hierarchical structure was investigated with two sets of data. Mean scores on the processes (levels of the taxonomy) were examined to determine if they decreased as the complexity of the levels increased. The intercorrelation matrix of process (level) scores were examined for the presence of simplical structure as described by Guttman's Radex Theory.

The second question or hypothesis about the transcendence of process was investigated by two sets of data. The intercorrelation matrix of process scores gathered over several contents were examined for the presence of circumplex structure as described by Guttman's Radex Theory. These intercorrelation matrices were also factor analyzed and the factors were examined for generality over content.

The third question, which dealt with the psychological structure of the taxonomy, was cast into an exploratory study. The exploratory study involved generating

orthogonal cognitive aptitude scores and determining their regression weights for each of the processes (levels) within and over contents.

Procedure

Tests. Special tests were constructed for the study because none was available commercially or from other investigators which would yield scores corresponding to the levels of the taxonomy.

Four taxonomy-type tests were constructed for use with ninth- through twelfth-grade students. Each of them consisted of six subtests corresponding to the major levels of the taxonomy. The tests were of the reading comprehension type in that a reading passage which presented the relevant content appeared in each test booklet and the test items were based on the content of it. The reading passages were selected on the basis of their probable interest value, probable ease of comprehension, and their unfamiliarity to students. The latter two specifications are important because it is necessary that content mastery of students be relatively equal so that score variability will reflect differential mastery of the cognitive processes. Of the four passages which were finally selected, two dealt with social science content, the Lisbon Earthquake and Stages of Economic Growth; and two dealt with science content, Atomic Structure and Glaciers.

Each test consisted of two parts. Part A included the Knowledge, Comprehension, Application, and Analysis items. Each of the subtests consisted of twenty four-choice items. Part B included the Synthesis and Evaluation items. There were five free-response Synthesis items and ten free-response Evaluation items. All subtests had a maximum possible score of twenty points. The administration time for each part is fifty to sixty minutes.

Preliminary forms of each of these taxonomy-type tests were pretested in a minimum of two versions. Revisions were based on observations of student performance during test administration, their reactions to structured questions at the conclusion of the test administration, the questions they asked during and after the test administration, and on statistical analyses of their responses.

The tests appear in Appendices A through D.

The cognitive aptitude tests were chosen from the Kit of Reference Tests for Cognitive Factors. Thirty-seven of them were selected for use in the study.

Test administration. Taxonomy-type tests were administered to ninth- through twelfth-grade students in ten Florida schools from five county school systems. Collectively the students represent a slightly above average group with respect to mental ability as determined by national norms. Tests were administered by project staff members with the assistance of local school personnel. Prior to testing at any school, orientation sessions were held for local teachers and administrators. The majority of students were administered all four taxonomy-type tests, which required eight hours of testing time. No student was administered fewer than two of the forms.

The thirty-seven cognitive factor tests were also administered to all students in one of the schools. Administering these required an additional eight hours from each student.

Approximately 1,600 students at each of the four grade-levels were administered the taxonomy-type tests, and approximately 275 of them from each grade level were administered the cognitive reference tests.

Test scoring. The forced-choice items of the Knowledge, Comprehension, Application, and Analysis items were scored electronically. The free-response Synthesis and Evaluation items were scored by a staff whose members were selected on the basis of qualifications resembling those of a classroom teacher, and then trained in the scoring procedures. This unusual qualification was specified to demonstrate that classroom teachers could score these items if called on to do so. Complete scoring instructions for each Synthesis and Evaluation item appear in Appendix E.

The cognitive aptitude tests were scored by a trained staff who had previous experience in scoring them.

Quality control checks were run intermittently on both scoring staffs. Data on rater reliability appear in Appendix E.

Analyses. To test the hypothesis of inverse relationship between mean performance and taxonomic level, mean scores for each subtest for each grade, nine through twelve, were calculated. Means were compared for each test form.

The hypothesis of simplical structure was investigated by techniques devised by Kaiser for scaling a simplex. It enables computing a statistic which can be used to select the order of variables in a matrix which will most closely approximate a perfect simplex. These analyses were completed for each form for each grade and for each form over all grades.

The hypothesis of transcendence of processes was tested by factor analysis and Guttman's circumplex analysis. Four 24 x 24 correlation matrices derived from the taxonomy-type subtests were factor analyzed. The unrotated matrices were interpreted because the available rotational techniques did not seem to be applicable to taxonomic data. Matrices consisting of single-process level intercorrelations taken over test forms were examined for circumplex structure. Analyses were calculated for each grade, nine through twelve, and for all grades combined.

The relationship between taxonomic level scores and cognitive aptitude scores was examined by a combination of factor analysis and multiple regression analysis. The 37 x 37 intercorrelation matrices of cognitive aptitude scores for each grade, separately, were factor analyzed. Ten factors were extracted and orthogonal cognitive aptitude scores were determined for each student. These factor scores were used to predict performance on each level of each taxonomy-type test. The multiple regression equations were examined for the presence of significant factors.

Results

The hypothesis of inverse relationship of mean performance and taxonomic level was generally supported. For the social science forms, means for all grades were in the predicted order. For the science forms there was an systematic reversal of means on the Synthesis and Evaluation subtests.

The simplex analyses offered some support to the hypothesis of hierarchical structure. For all analyses of the social science forms for grades ten, eleven,

twelve, and all grades combined, the hypothesized order was the best order in terms of simplex order. For one science form at all grade levels, the same order of variables led to simplical structure and it was different from the hypothesized order. It placed Evaluation between Knowledge and Comprehension.

On the basis of these results, the conclusion was drawn that there was a clear tendency for the empirical data to support the imputed hierarchical structure of the taxonomy.

The hypothesis of the generality of process was investigated by circumplex analysis and factor analysis. The former analyses lent support to the hypothesis although no perfect circumplexes were obtained. Closest approximations to the circumplex occurred with the subtests on Knowledge and Evaluation. The factor analysis tended to corroborate the results of the circumplex analysis in that factors associated with Knowledge, Synthesis, and Evaluation appeared to be common over forms and grades. However, the usefulness of the factor analysis results were seriously impaired because unrotated factor matrices had to be interpreted and the majority of the factors which were extracted were mixtures of content and process. The presence of these mixed factors tends to discredit the hypothesis and to suggest that the scores are determined by highly complex interactions of content and process.

The studies to determine the aptitude structure of the taxonomic levels were impaired because only eight common factors (six were common to four grades and two were common to three grades) could be extracted, whereas sixteen were expected to emerge, and because the majority of factors from the analysis of the taxonomy-type tests were mixtures of content and process. Nevertheless, certain patterns did emerge which suggested somewhat systematic changes in factor structure over process levels and grades. However, no claim was made about them because of the obvious need for more refined analytic techniques and data from more refined taxonomy-type tests, and because of the mixed nature of the taxonomy factors which were extracted.

B. Recommendations

The recommendations which are presented here arose from problems which were encountered during the study. Many of them were not anticipated, none of them was

satisfacotrily resolved, and it seems that all of them deserve concerted study prior to executing further investigations of the construct validity of the taxonomy. The problems are listed here and are accompanied by brief explanations. Extended discussions of them appear in the body of the report.

Knowledge as a Process

Some doubt exists about whether Knowledge, the lowest level of the taxonomy, should be regarded as the process of recall or a measure of stored content. Perhaps this question arises because the measurement of the process of recall seems to be indivisible from the measurement of stored content. The problem is fundamentally important because validation studies of the taxonomy are contingent on gathering data about the execution of processes by students all of whom have complete mastery of the relevant content. When a student fails a Knowledge item, a perplexing dilemma arises. Was the process deficient or had the content not been stored? Presently available measurement techniques do not enable discriminating one from the other. The consequence of not having this discriminative measurement partially invalidates the measurement of the higher level processes because the failure on the higher-level item cannot be attributed exclusively to inadequate mastery of the relevant process. In other words, process score variability might contain a component of variability in content knowledge. Its presence decreases the validity of the process data and impedes efforts to validate the structure of the taxonomy.

Until more refined measurement techniques are available it would probably be sensible to cause students to overlearn the relevant content to the extent that the investigator would have little doubt that the content has been stored. But even when that precaution is taken, the retrieval from storage (Knowledge) will be a function of the item's stimulus value in triggering the retrieval. We have prepared many items all dealing with the same bit of content but each differing in the cues given in the stems and have found that item difficulties vary substantially and predictably over a group which has had equal exposure to the content.

Item Analysis Data

Guidelines for the legitimate interpretation of item analysis data from taxonomy-type tests must be established. When dealing with conventional tests one usually strives to maximize the number of discriminations by selecting items at the mid-range of difficulty and, if several subtests are involved, to maximize the homogeneity of each and to minimize subtests intercorrelations. These procedures do not seem to be applicable when using item data as a step in test refinement.

A taxonomy-type test will undoubtedly contain subtests (or a mass of items which are scored according to process keys) which correspond to the levels (Processes) of the taxonomy. According to the claimed structure of the taxonomy, these subtests will vary systematically in complexity. Consequently, one can probably expect that as level increases then mean item difficulties will decrease. Furthermore, subtests will be intercorrelated and these coefficients should decrease systematically according to the relative remoteness in the taxonomic structure of the subtests which are correlated.

Therefore, it seems sensible to devise theoretically a set of expectations about the desirable range of item difficulties for each subtest and the pattern of item-subtest correlations which an item should exhibit over all subtests.

If one unwittingly applied conventional item analysis and selection techniques then he will certainly violate the structure of the taxonomy and produce a test which probably will yield invalid data.

Triangular Bivariate Distributions

Assuming the structure of the taxonomy, then it probably follows that bivariate distributions among subtests will be triangular. The lower level subtest will establish the ceiling for scores on the higher level subtest. The majority of bivariate distribution which we plotted from our data exhibited this characteristic. Its presence raises questions about what measure of association can be calculated legitimately from such data. We acknowledge that our data had this characteristic; nevertheless we calculated product-moment correlations because we could not identify any measure of association that could be applied legitimately to these data. We are

unaware of the gravity of applying an inappropriate statistic, perhaps the development and employment of legitimate techniques of measuring association will eventually lead to discrediting the results which are presented in this report.

The comments in this section and the preceding one on item analysis data are based on the assumption that taxonomy-type tests will contain the relevant content on which the items are based. Under this condition, means will probably decrease as item levels increase and pairs of subtests will have bivariate distributions. In short, complexity and difficulty will probably be similar. But if content is not supplied with the test items, then Guttman's differentiation of complexity and difficulty must be taken into account and it is likely that our suggestions and data will be irrelevant.

Intended and Actual Processes

The taxonomists clearly stated that the phenomena to be classified by the taxonomy are the processes which test items are intended to evoke. Our study is based on the responses which items actually evoked. Consequently, if one takes a very narrow view of the validation process, then our study does not deal with the validation of the taxonomy. However, we believe that if the taxonomy is used as a guide in structuring educational experiences then it must be shown to be descriptive of empirical phenomena or corroborated by them.

The more important question seems to be about the tenability of the assumption on which our tests were constructed; specifically, that a correct response to an item implies use and mastery of the appropriate process and that an incorrect response indicates the contrary. We believe that if the logistical problems were of no consequence then the most pertinent data which could be gathered for validating the taxonomy would be the recordings of oral problem solving processes by students. We definitely regard our data as second best from the methodological standpoint but perhaps the most relevant which could be collected from masses of students when judged from the practical standpoint.

The Evaluation Process

The Evaluation process as described in the taxonomy is one which deserves further study. First, Evaluation in terms of external criteria, one of the two sublevels, seems to introduce content which is alien to the content on which the test items are based. Second, the studies by McGuire suggest that Evaluation according to internal criteria is a process which is less complex than is Synthesis. We found it to be the most difficult process for which to write items and our analyses reveal that it was the one which most frequently did not conform to the hypothesized order of the levels.

Norms for Taxonomy-type Tests

If taxonomy-type tests are constructed and come into wide usage, then it will probably be necessary to devise some meaningful derived score for the scores which the test yields. The problem is quite complex and probably cannot be solved by current norming techniques. The derived score will have to be created with due regard to the following points. First, a taxonomy-type test will probably consist of a series of contingent tests in which the attained score on a lower level test establishes the maximum score on the test of the next higher level. Second, since performance will decrease as level of subtest increases, the raw score distributions will become increasingly skewed and on the higher levels the frequency of "zero" scores will probably be substantial. Third, the raw scores seem to have both ipsative and normative characteristics. The score distribution on any taxonomy subtest (which will be a power test) represents the distribution of the persons over that trait. On the other hand, a person's scores over the subtests are not independent because a given subtest score probably establishes the ceiling on the score which might be attained on the subtest which appears next above in the hierarchy.

Item Types

We attempted to prepare items in the multiple-choice format for all levels of the taxonomy. We could not write multiple-choice Synthesis and Evaluation items which were regarded by the project staff as likely to evoke the intended process. The staff seemed to classify as Analysis multiple-choice items which were written for the Evaluation

process. Multiple-choice items which were written for Synthesis were frequently classified as Analysis or even lower in the structure. As a consequence, we adopted the free-response format for Evaluation and Synthesis. The scoring time required for responses to these items was sufficiently great to infer that taxonomy-type tests which include free-response items will not be as feasible for mass administration as are the typical achievement tests, the items of which seem to be confined to the three lowest levels of the taxonomy. Because the Synthesis and Evaluation processes are critically important and because of the nearly predictable unacceptance of survey instruments which include free-response items, it seems especially appropriate to conduct research on items writing techniques and formats which might make these processes amenable to multiple-choice assessment.

Pyramid Tests

If a student fails to answer correctly all Knowledge items which appear in that subtest, then the scores he attains on the higher level subtests are of doubtful validity because one does not know whether to attribute these subsequent failures to lack of stored content or to inability to recall information which is stored. Therefore, it appears that a student's mastery of the processes can be assessed only when the test items employ content which he has available. It seems theoretically possible to prepare sizeable numbers of items at each of the process levels and all based on the same universe of content. If these items could be appropriately identified according to the level and the bits of content which they require, then a student could first be exposed to the Knowledge items and on the basis of his responses to them, then one could determine the precise elements of content which he has available for use. On the basis of this information, items could immediately be selected for all succeeding levels which would not require content knowledge which he does not possess; thus, the notion of a pyramid test. Undoubtedly, each test would be unique, being patterned especially for the student. The logistics of the technique would require that all items and their characteristics be placed in a computer storage and that the student interact with the computer through an input-output device or teaching terminal, and that the test be assembled on the basis of the student's responses to the first set of items which is presented to him.

PRECEDING PAGE BLANK-NOT FILMED

REFERENCES

Anderson, June S. A comparative study of chemical educational material study and traditional chemistry in terms of students' ability to use selected cognitive processes. Unpublished doctoral dissertation, Florida State University, 1964.

Bloom, B. S. (ed.) Taxonomy of educational objectives, handbook I: cognitive domain. New York: David McKay and Co., 1956.

Crawford, W. R. A validation of the structure and generality of a Taxonomy of intellectual processes. Unpublished doctoral dissertation, Florida State University, 1966.

Dunn, S. S. & Turner, M. ACER Prematriculation physics examination, report no. 2. Victoria: Australian Council for Educational Research (1959).

French, J. W., Ekstrom, Ruth B., & Price, L. A. Manual for Kit of reference tests for cognitive factors. Princeton: Educational Testing Service, 1963.

Guttman, L. A new approach to factor analysis: the radex. In Lazarsfeld, P. (ed.), Mathematical thinking in the social sciences. Glencoe, Illinois: The Free Press, 1964.

Harman, H. H. Modern factor analysis. Chicago: The University of Chicago Press, 1960.

Hurley, J. R. & Cattell, R. B. The procrustes program: producing direct rotation to test a hypothesized factor structure. Behavioral Science, 1962, 7, 258-262.

Kaiser, H. F. Scaling a simplex. Psychometrika, 1962, 27 (2), 155-162.

Kellogg, R. E. American history test bank, secondary level. San Diego, California: San Diego County Department of Education, 1964.

McFall, R. W. & Chambers, W. M. Unpublished research report. University of Kentucky.

McGuire, Christine. Research in the process approach to the construction and analysis of medical examinations. Paper read at National Council on Measurement in Education, Chicago, Feb., 1963.

Saunders, D. R. "A computer program to find the best-fitting orthogonal factors for a given hypothesis." Psychometrika, 1960, 25 (2), 199-205.

Smith, A. B. An analysis of the scalability of the "Knowledge" and "Comprehension" levels of the Taxonomy of educational objectives: cognitive domain. Paper read at American Educational Research Association, Chicago, Feb., 1965.

Stanley, J. C. & Bolton, D. T. Book Reviews. Psychol. Measmt., 1957, 17, 631-634.

Stoker, H. W. & Kropp, R. P. Measurement of cognitive processes. J. Educ. Measmt., 1964, 1, 39-42.

Thomas, Alice M. Levels of cognitive behavior measured in a controlled teaching situation. Unpublished masters thesis. Cornell University, 1965.

Tyler, Leona E. The psychology of human differences. (3rd ed.). New York: Appleton-Century-Crofts, 1965, Chap. 9, pp. 211-235.

APPENDIX A

Table of Contents

ATOMIC STRUCTURE

Directions

This is a test of your ability to apply skills and understandings which you have been learning since you entered school. You will be asked to read a passage called "Atomic Structure" and then you will be asked to answer questions about it. You may refer to the passage any time during the test.

Your answer must be marked on the answer sheet which has been provided. Do not make any stray marks. If you make an error, erase it completely before marking your new answer.

Make no marks in the test booklet.

The following is a sample question to show you how your answers are to be marked. Study the sample and if you have any questions raise your hand.

SAMPLE ON SEPARATE ANSWER SHEET

1. The article you will read is about
   1. atomic structure
   2. northeast floods
   3. last days of World War II
   4. the Spanish-American War

1. 1 ▬ 2 ___ 3 ___ 4 ___ 5 ___

You should answer as many questions as you can. Do not spend a great amount of time on any one question. If you are in doubt about the answer to a question, then guess. Remember that you may refer to the reading passage at any time during the test.

R.P. Kropp and H.W. Stoker, Editors
Department of Educational Research and Testing
and Institute of Human Learning
Florida State University
Tallahassee

April, 1965

Experimental edition; to be used for research purposes.

This experimental test was constructed as part of a research project supported by the U. S. Office of Education under Contract Number OE-4-10-019, Project Number 2117.

The copyright holders of the materials which appear in the reading passage of the test kindly permitted the project staff to use the materials experimentally and the holders subsequently extended this permission to all persons who wish to use the experimental version of the test for research purposes.

ATOMIC STRUCTURE

Dmitri Mendeleef now arranged the elements into seven groups, starting with lithium (at. wt. 7), and followed by beryllium (at. wt. 9), boron (11), carbon (12), Nitrogen (14), oxygen (16) and fluorine (19). The next element, in the order of increasing atomic weight, was sodium (23). This element resembled lithium very closely in both physical and chemical properties. He, therefore, placed it below lithium in his table. After placing five more elements, he came to chlorine, which had properties very similar to fluorine, under which it miraculously fell in his list. In this way he continued to arrange the remainder of the elements. When his list was completed, he noticed a most remarkable order. How beautifully the elements fitted into their places. The very active metals lithium, sodium, potassium, rubidium and cesium fell into one group (No. I). The extremely active non-metals, fluorine, chlorine, bromine and iodine all appeared in the seventh group.

Mendeleef had discovered that the properties of the elements "were periodic functions of their atomic weights," that is, their properties repeated themselves periodically after each seven elements. What a simple law he had discovered! But here was another astonishing fact. All the atoms in Group I united with oxygen, two atoms to one. All the atoms of the second group united with oxygen, atom for atom. The elements in Group III joined with oxygen, two atoms to three. Similar uniformities prevailed in the remaining groups of elements. What in the realm of nature could be more simple? To know the properties of one element of a certain group was to know, in a general way, the properties of all the elements in that group. What a saving of time and effort for his chemistry students!

Could his table be nothing but a strange coincidence? Mendeleef wondered. He studied the properties of even the rarest of the elements. He re-searched the chemical literature lest he had, in the ardor of his work, misplaced an element to fit in with his beautiful edifice. Yes, here was a mistake! He had misplaced iodine, whose atomic weight was recorded at 127, and tellurium, 128, to agree with his scheme of things. Mendeleef looked at his Periodic Table of the Elements and said that it was good. With the courage of a prophet, he made bold to say that the atomic weight of tellurium was wrong; that it must be between 123 and 126 and not 128, as its discoverer had determined. Here was downright heresy, but Dmitri was not afraid to buck the established

order of things. For the present, he placed the element tellurium in its proper position, but with its false atomic weight. Years later his action was upheld, for further chemical discoveries proved his position of tellurium to be correct. This was one of the most magnificient prognostications in chemical history.

Perhaps Mendeleef's table was now free from flaws. Again, he examined it, and once more he detected an apparent contradiction. Here was gold with the accepted atomic weight of 196.2 placed in a space which rightfully belonged to platinum, whose established atomic weight was 196.7. The fault-finders got busy. They pointed out this discrepancy with scorn. Mendeleef made brave enough to claim that the figures of the analyst, and not his table, were inaccurate. He told them to wait. He would be vindicated. And again the balance of the chemist came to the aid of the philosopher, for the then-accepted weights were wrong and Mendeleef was again right. Gold had an atomic weight greater than platinum. This table of the queer Russian was almost uncanny in its accuracy!

Mendeleef was still to strike his greatest bolt. Here were places in his table which were vacant. Were they always to remain empty, or had the efforts of man failed as yet to uncover some missing elements which belonged in these spaces? A less intrepid person would have shrunk from the conclusion that this Russian drew. Not this Tartar, who would not cut his hair even to please his Majesty, Czar Alexander III.. He was convinced of the truth of this great generalization, and did not fear the blind, chemical skeptics.

Here in Group III was a gap between calcium and titanium. Since it occurred under boron, the missing element must resemble boron. This was his eka-boron which he predicted. There was another gap in the same group under aluminum. This element must resemble aluminum, so he called it eka-aluminum. And, finally he found vacant space between arsenic and eka-aluminum which appeared in the fourth group. Since its position was below the element silicon, he called it eka-silicon. Thus he predicted three undiscovered elements, and left it to his chemical contemporaries to verify his prophecies. Not such remarkable guesses after all--at least not to the genius Mendeleef!

Mendeleef had made places for more than sixty elements in his Table. Three more he had predicted. What of the other missing building blocks of the universe? Twenty-five years after the publication of Mendeleef's Table, two Englishmen, following a clue of Cavendish, came upon a new group of elements of which even the Russian had never dreamed.

These elements constituted a queer company--the Zero Group, as it was later named. Its members, six in number, are the most unsociable of all the elements. Even with that ideal mixer, potassium, they will not unite. Fluorine, most violent of all the non-metals cannot shake these hermit elements out of their inertness. Moissan tried sparking them with fluorine, but failed to make them combine. Besides, they are all gases, invisible and odorless. Small wonder they had remained so long hidden.

Besides these six Zero elements, seventeen other elements were unearthed, so that, a year after Mendeleef died in 1907, eighty-six elements were listed in the Periodic Table. Table 1 shows Mendeleef's table in 1870.

Two years after Mendeleef was laid beside the grave of his mother and son, Pattison Muir declared that "the future will decide whether the Periodic Law is the long looked for goal, or only a stage in the journey; a resting place while material is gathered for the next advance."

History has shown that the Periodic Law weathered time beautifully. Throughout the years the current Periodic Table has had many similarities to Mendeleef's table. Table 2, the current periodic table, reflects one major change. For many years the atomic weights of elements have been based on the weight of oxygen - 16.000; this table (Table 2) is based on the weights of carbon - 12.000.

Current Periodic Tables, such as Table 2, are built on the concept of atomic number rather than atomic weight. The present Periodic Law can be stated thusly: The properties of the elements are periodic functions of the atomic numbers. Table 2 makes use of this concept and also uses the concept of increasing energy. To illustrate: In Table 2, oxygen (O) - at. wt. = 16.0; at. no. = 8; Group =VI. We now see that every atom has associated with it a characteristic atomic number and a characteristic atomic weight.

The atomic weight of an atom is the total weight of its parts which are protons, electrons, and neutrons. Each proton has a weight of one (1); each electron has a weight of zero (0); and each neutron has a weight of one (1). The reason the table based on atomic weight and the table based on atomic number are similar is that the atomic number is equal to the number of protons in an atom.

In addition to weight, the parts of an atom have characteristic electrical charges. Protons are positive (+), electrons are negative (-), and neutrons are neutral (0). In every element there are equal numbers of protons and electrons. There are situations when an atom does not have equal numbers of electrons and protons; when this occurs, the atom is called an ion.

Characteristics of the Particles of an Atom

Particle	Weight	Charge
Proton	1	+1
Electron	0	-1
Neutron	1	0

Adapted from

Bernard Jaffe, "Mendeleef" from The World of Mathematics by James R. Newman; Vol. 2, Simon and Schuster, New York, 1956, pp. 919-931.

Table 1. Mendeleef 1870

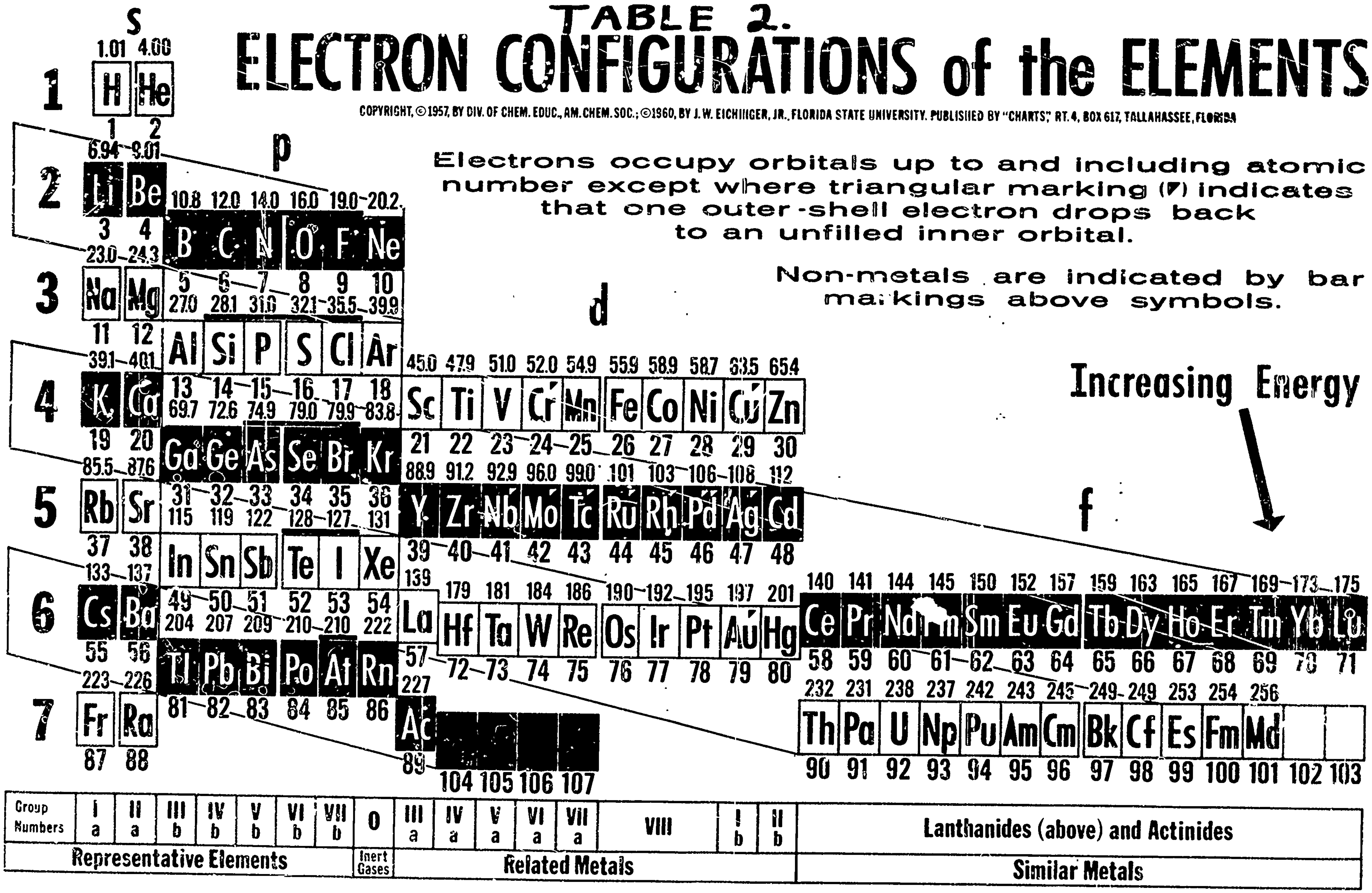
Groups:		I	II	III	IV	V	VI	VII	VIII
Typical Elements:		H Li	Be	B	C	N	O	F	
Periods 1.	Series 1. 2.	Na K	Mg Ca	Al -	Si Ti	P V	S Cr	Cl Mn	FeCoNiCu
2.	3. 4.	(Cu) Rb	Zn Sr	- Y?	- - Zr	As Nb	Se Mo	Br -	RuRhPdAg
3.	5. 6.	(Ag) Cs	Cd Ba	In -	Sn Ce	Sb -	Te -	I -	- - -
4.	7. 8.	- -	- -	- -	- -	- Ta	- -	- -	OsIrPtAu
5.	9. 10.	(Au) -	Hg -	Tl -	Pb Th	Bi -	- U	- -	- - -

TABLE 2.

ELECTRON CONFIGURATIONS of the ELEMENTS

Electrons occupy orbitals up to and including atomic number except where triangular marking (◤) indicates that one outer-shell electron drops back to an unfilled inner orbital.

Non-metals are indicated by bar markings above symbols.



1. Mendeleef placed the chemical element sodium (Na) in group
   (K) 1. I
   2. II
   3. III
   4. IV

2. Which pair of elements is in the same periodic group?
   (K) 1. Sodium, cesium
   2. Sodium, chlorine
   3. Bromine, rubidium
   4. Oxygen, carbon

3. Mendeleef discovered that the properties of elements
   (K) 1. were periodic functions of atomic numbers.
   2. were periodic functions of their atomic weights.
   3. were dependent upon temperature and pressure.
   4. could not be ordered.

4. Atoms of which group unite with atoms of oxygen in the ratio of 2 to 3?
   (K) 1. I
   2. II
   3. III
   4. IV

5. All elements in the same group have similar
   (K) 1. atomic weights.
   2. general properties.
   3. atomic numbers.
   4. specific properties.

6. How many elements are in the Zero Group?
   (K) 1. 3
   2. 6
   3. 9
   4. 17

7. Which element is known as the ideal mixer?
   (K) 1. Hydrogen (H, #1)
   2. Oxygen (O, #8)
   3. Flourine (F, #9)
   4. Potassium (K, #19)

NOTE: The symbol (#) is used in several questions in this test. This symbol stands for atomic number.

8. According to the passage, elements that do not combine with any other elements belong to the
(K) 1. Second Group.
2. Gas Group
3. Zero Group.
4. Metal Group.

9. The Zero Group elements might be described as
(K) 1. elements with atomic weights of zero.
2. highly active gases.
3. highly active non-metals.
4. invisible and odorless gases.

10. The chemical world has had the advantage of a Periodic Table for approximately how many years?
(K) 1. 50
2. 100
3. 150
4. 200

11. An atom with unequal numbers of protons and electrons is
(K) 1. a neutron.
2. an electron.
3. an element.
4. an ion.

12. In Table 2, which of the following symbols is used to indicate an element in which one outer shell electron drops back to an unfilled inner orbital?
(K) 1. d
2. triangle
3. black square
4. Roman numeral

13. In Table 2, which of the following symbols is used to indicate a non-metal?
(K) 1. p
2. arrow
3. bar
4. white square

14. Neon (Ne) has an atomic number of
(K) 1. 10
2. 11
3. 12
4. 16

15. Cobalt (Co) has an atomic weight of
(K) 1. 58.9
2. 52.0
3. 83.8
4. 101.0

16. In Table 2, rhenium (Re, #75) is a member of Group
(K) 1. III
2. IV
3. VI
4. VII

17. Iridium (Ir) has an atomic weight of
(K) 1. 186
2. 190
3. 192
4. 195

18. In Table 2, astatine (At) is a member of Group
(K) 1. III
2. IV
3. V
4. VII

19. In Table 2, most of the inert gases appear in Group
(K) 1. 0
2. I
3. II
4. VI

20. Table 2 indicates that sulphur (S, #16) is
(K) 1. a yellow powder.
2. a non-metal.
3. a typical element.
4. inert.

21. Mendeleef thought he had misplaced iodine and tellurium in his table because
(C) 1. the accepted weight of iodine was less than that of tellurium.
2. their known properties were not the same.
3. their known properties were the same.
4. the accepted weight of iodine was more than that of tellurium.

22. According to Table 2, the number of electrons occupying orbitals in the element aluminum (Al, #13) is
(C)
1. 5
2. 12
3. 13
4. 27

23. What is the relationship between
(An)
A. Group number, and
B. Atomic number
1. An increase in A is usually accompanied by an increase in B.
2. An increase in A is usually accompanied by a decrease in B.
3. An increase in A is usually accompanied by either an increase or decrease in B.
4. A is determined by B.

24. Which of the following best describes Mendeleef's attitude toward his findings?
(Ap)
1. He would apply his theory regardless of the consequences to his reputation.
2. New fields of research were open to chemists.
3. The relationships he had discovered were possibly coincidental.
4. The results would probably hold true even though his work was largely guesswork.

25. The correct ordering of elements according to <u>increasing</u> atomic weights is
(C)
1. F, Cl, Na.
2. Na, Cl, F.
3. F, Na, Cl.
4. Cl, F, Na.

26. Which of the following is the best estimate of the year in which Mendeleef was born?
(An)
1. 1800
2. 1810
3. 1840
4. 1860

27. Mendeleef's table can be best described as
(C)
1. a chemical table arranged in increasing atomic numbers.
2. a table of chemical elements arranged in a precise order.
3. an orderly grouping of chemical elements.
4. a periodic table of chemical elements.

28. The symbol for the element which has only one proton and and one electron is
(Ap) 1. H (#1)
2. He (#2)
3. Y (#39)
4. Ac (#89)

29. According to Table 2, which of the following elements has the greatest amount of chemical energy?
(C) 1. Ca (#20)
2. Kr (#36)
3. Ag (#47)
4. Tm (#69)

30. Group II in Mendeleef's table which contains magnesium (Mg), calcium (Ca) and strontium (Sr), has spaces for other elements. Which one of the missing elements is shown in that group in Table 2?
(Ap) 1. Barium (Ba)
2. Radium (Ra)
3. Cesium (Cs)
4. Zinc (Zn)

31. Identify the element with the following characteristics:
(Ap) A. Its atomic weight is between 133 and 256.
B. It is a representative element.
C. It is a non-metal.
This element is
1. Te (#52)
2. I (#53)
3. Po (#84)
4. At (#85)

32. How many electrons does the element nitrogen (N) have?
(Ap) 1. 0
2. 5
3. 7
4. 14

NOTE: The following information must be used in answering questions 33 - 35. Four mystery elements with atomic weight and number are given below:

Element	Atomic Weight	Atomic Number
Xa	352	190
Xb	351	191
Xc	357	193
Xd	368	189

33. The correct order according to Mendeleef (Table 1) is
(An) 1. , Xa, Xb, Xc, Xd
2. Xb, Xa, Xc, Xd
3. Xd, Xc, Xb, Xa
4. Xd, Xa, Xb, Xc

34. The correct order according to Table 2 is
(An) 1. Xa, Xb, Xc, Xd
2. Xc, Xb, Xa, Xd
3. Xb, Xa, Xc, Xd
4. Xd, Xa, Xb, Xc

35. Which element would be placed in the same relative position in both tables?
(An) 1. Xa
2. Xb
3. Xc
4. Xd

* * * * *

36. Why did Mendeleef place elements Cu, Ag, and Au in both Groups I and VIII in his table?
(An) 1. They have characteristics similar to elements in both groups.
2. He felt that undiscovered elements similar to these would be placed in these spaces in Group I.
3. He thought he had not placed enough elements in Group I.
4. He felt that all elements could be classified in seven groups.

37. According to the reading passage the Zero Group
(C) 1. cannot be scientifically isolated.
2. will not unite with atoms of other elements.
3. do not fulfill the conditions of the Periodic Table.
4. Are the only elements in gaseous form.

38. Very active metals occur in Group I and very active non-metals occur in Group VII. A prediction about elements in Group IV would be that they
(An) 1. have only non-metallic properties.
2. have only metallic properties.
3. do not resemble metals or non-metals in their properties.
4. resemble both metals and non-metals in their properties.

39. If a new element were discovered with an atomic weight of 258, which of the following is the best estimate of its atomic number?
(Ap) 1. 102
2. 106
3. 107
4. Must be larger than 107.

40. According to Table 2, calcium (Ca, #20) has chemical properties most similar to
(C) 1. Potassium (K, #19)
2. Aluminum (Al, #13)
3. Scandium (Sc, #21)
4. Zinc (Zn, #30)

41. Why did Mendeleef leave gaps in his table?
(Ap) 1. He believed that forthcoming discoveries would need the spaces.
2. He could not find elements with the characteristics for these spaces.
3. Atomic weights of known elements could not be determined.
4. He did not know enough about the chemistry of known elements.

42. According to Table 1, the chemical properties of which pair of elements are most similar?
(C) 1. Rb and Hg
2. Mg and Al
3. Ni and Cu
4. Na and K

NOTE: The symbol (#) stands for atomic number.

43. How many protons does the element Ac have?
(Ap) 1. 0
2. 56
3. 89
4. 138

44. The ratio of neutrons to protons in the element carbon (C, #6) is
(Ap) 1. 1 to 1
2. 1 to 2
3. 4 to 3
4. 2 to 1

45. According to Table 2, which of the following is a characteristic of hydrogen (H, #1)?
(Ap) 1. One outer shell electron drops back.
2. It is a non-metal.
3. It is heavier than He.
4. It is a metal.

46. If their existence had been known, which of the following pairs of elements would have been most difficult for Mendeleef to place?
(Ap) 1. Am (#95) and Cm (#96)
2. Cm (#96) and Bk (#97)
3. Bk (#97) and Cf (#98)
4. Tb (#65) and Bk (#97)

47. Which pair of elements is most likely to combine?
(C) 1. Ne (#10) and Ar (#18)
2. K (#19) and Cl (#17)
3. Xe (#54) and K (#19)
4. Rn (#86) and Cs (#55)

48. Atoms of magnesium (mg, #12) and sulphur (S, #16) combine in a ratio of
(An) 1. 1 to 3
2. 1 to 1
3. 2 to 1
4. 3 to 2

49. The atomic weight of oxygen (O) when compared to that of flourine (F) is
(C) 1. lower
2. the same
3. not comparable
4. higher

50. According to Mendeleef in Table 1, the number of undiscovered elements in Group II was
(C) 1. 1
2. 2
3. 3
4. 4

51. Which of the following is most active?
(C) 1. Fluorine (F)
2. Chlorine (Cl)
3. Bromine (Br)
4. Iodine (I)

52. Which of the following non-metals is most active?
(Ap) 1. Oxygen (O, #16)
2. Chlorine (Cl, #17)
3. Iodine (I, #52)
4. Actinium (Ac, #89)

53. How many neutrons does the element helium (He) have?
(Ap) 1. 0
2. 2
3. 3
4. 4

54. The element which has 9 electrons, 10 neutrons, and 9 protons is
(Ap) 1. Be (#4)
2. F (#9)
3. Ni (#28)
4. Ne (#10)

55. According to Table 2, the chemical elements americium (Am, #95) and curium (Cm, #96) are
(C) 1. representative elements.
2. inert gases.
3. related metals.
4. similar metals.

56. An ion which has a charge of -1
(Ap) 1. has an extra proton.
2. is missing an electron.
3. has an extra electron.
4. has an extra proton <u>and</u> electron.

57. An atom with unequal numbers of protons and neutroms is
(Ap) 1. Hydrogen (H, #1)
2. Helium (He, #2)
3. Carbon (C, #6)
4. non-existent

58. The element Mendeleef called eka-silicon (eka-Si) is now known as
(C) 1. gallium (Ga, #31)
2. germanium (Ge, #32)
3. zirconium (Zr, #40)
4. Niobium (Nb, #41)

59. An ionic form of an element has a total of 18 electrons and a charge of -1. The element's atomic number is
(An) 1. 16
2. 17
3. 18
4. 19

<u>NOTE:</u> The following information applies to questions 60-62: When a chemist writes the formula for a compound, he shows with subscripts the number of atoms of each element in a "molecule" of the compound.

60. The formula for aluminum oxide (Al, #13 and O, #8) is
(An) 1. Al_3O_2
2. AlO_3
3. Al_2O_3
4. Al_2O

61. Calcium (Ca, #20) and tellurium (Te, #52) combine to form
(An) 1. CaTe
2. $CaTe_2$
3. Ca_2Te_3
4. Ca_3Te_2

62. Suppose one was given a compound whose formula was "?"$_2$O. Which group would contain the element indicated by the question mark?
(Ap) 1. I
2. II
3. III
4. IV

* * * * * * *

63. If given the atomic number of an element, the atomic number of the next element in that group may always be found by
(An) 1. adding 8 to the known atomic number.
2. adding 1 if in Group I, 2 if in Group II, 3 if in Group III, etc.
3. multiplying the known atomic number by 2.
4. none of the above.

64. Given an element of the following characteristics:
(An) A. atomic weight greater than Si.
B. atomic weight less than V.
C. general characteristics similar to N.
This element is
1. Cv
2. P
3. Nb
4. Ti

65. After studying his original Table, Mendeleef decided that
(C) 1. iodine should be placed in front of tellurium.
2. iodine had been assigned a false atomic weight.
3. tellurium had always been in its proper position.
4. tellurium should be placed in front of iodine.

66. Since the element names eka-silicon (eka-Si) by Mendeleef combines with oxygen, the number of atoms of eka-silicon needed to combine with two atoms of oxygen is
(An) 1. 1
2. 2
3. 3
4. 4

67. Which of the following metals has general characteristics most similar to those of non-metals?
(An) 1. Na (#11)
2. K (#19)
3. Te (#52)
4. Po (#84)

68. According to Table 1, an element is missing between Nb and Ta in Group V. This element is
(An) 1. one of the Actinides shown in Table 2.
2. shown by the unknown element with atomic number 109 in Table 2.
3. most likely in Group I in Table 2.
4. not shown in Group V in Table 2.

69. According to Mendeleef, the atomic weight of tellurium (Te, #52) should be between 123 and 126, instead of 128. Which of the following is the best explanation for the placement of Te in Table 2?
(An) 1. The atomic weight of Te was wrong as Mendeleef predicted.
2. Table 2 is based on atomic number.
3. The atomic weight of iodine (I, #53) proved to be greater than that of Te.
4. The established atomic weight of Te was correct, and his placement was incorrect.

70. When oxygen combines with an element from Group I, the ratio of Group I atoms to oxygen atoms is
(C) 1. two to three.
2. two to one.
3. one to one.
4. three to two.

71. Which of the following best explains why Mendeleef did not leave spaces for the Zero Group elements in his table?
(An) 1. Their physical characteristics differ from those of the Typical Elements.
2. They cannot be classified as metals or non-metals.
3. He could not predict their atomic weights because they are invisible.
4. A table based on atomic weights would not indicate omission of an entire group.

72. The heaviest non-metal is
(C) 1. I (#53)
2. Te (#52)
3. Po (#84)
4. At (#85)

73. According to Table 2,
(C) 1. H (#1) is heavier than He (#2).
2. La (#57) is heavier than Ce (#58).
3. Th (#90) is heavier than Pa (#91).
4. Po (#84) is heavier than At (#85).

74. According to Mendeleef's table there is an element missing below Ta in Group V. Table 2 indicates that this element

(An) 1. has not been discovered.
2. has an atomic weight of approximately 229.
3. has an atomic number of 107.
4. all of the above.

75. An important use for the periodic table would be to

(C) 1. develop new methods of classifying elements.
2. give the total number of elements.
3. predict new elements.
4. predict properties of new elements.

76. The atomic weight of water (H_2O) is

(Ap) 1. 3
2. 10
3. 18
4. 20

77. Approximately what per cent of the weight of sodium (Na, #11) is neutrons?

(Ap) 1. 30
2. 50
3. 70
4. 90

78. Suppose an element were discovered which united with oxygen two atoms to one. The best estimate of the element's atomic number is

(An) 1. 87
2. 107
3. 111
4. 267

79. Mendeleef's basis for placing tellurium (Te #52) before iodine (I, #53) was

(C) 1. the general characteristics of the two elements indicated that their order should be reversed.
2. the discoverer of tellurium was noted for careless work, therefore mendeleef assumed him to be in error.
3. sould experimentation proved that the weight of iodine was wrong.
4. the atomic numbers of the two elements indicated that their order should be reversed.

80. Which of the following was an assumption basic to Mendeleef's work?

(Ap)
1. All elements could be classified in 7 groups.
2. Miscalculated atomic weights would show up in his table.
3. Elements can be arranged in an orderly sequence based on chemical properties.
4. Many elements were yet to be discovered; therefore, his table would be incomplete.

ATOMIC STRUCTURE
(Booklet Number 2)

Directions

Your answers to each of items 81-95 should be written directly on the test booklet in the space provided below each item. Should you need additional space for an item, write "continued"; then write the remainder of your answer on the back of the last page of the booklet.

You received with this test another booklet which contains the article about Atomic Structure. You may refer to it any time during the examination. You might wish to scan the article after you read these directions to refresh your memory.

Your answers to these items will depend on your understanding and thoughts about the article. Each of you might write an entirely different answer to an item, because there are many possible answers to each. So please write whatever you believe to be appropriate.

Answers to these items do not appear as such in the article. Therefore, you will certainly answer incorrectly if you copy material directly from the article.

81. (Sy) A chemist claims that he was able to make a sample of xenon, a member of the Zero Group, combine with fluorine. Briefly explain what might have happened to produce this result.

82. (Sy) The article states that Mendeleef started his table by arranging the elements into seven groups. How might his results have differed if he had started by arranging the elements into seventeen groups?

83. (Sy) Mendeleef predicted elements to fill three of the gaps in his table, yet he failed to predict elements to fill these gaps.

84. (Sy) Write a brief history of the development of the periodic tables. The article will provide a basis for this history, but you should supply events which could have occurred in addition to those given.

85. (Sy) Suppose that you were the first chemist to visit another planet. What steps would you take in constructing a periodic table of the elements found on that planet?

86.-95. Suppose that you have been asked to evaluate the author's qualifications to write an article on the development of the Periodic Table. Listed below are several statements which could be true about the author. Assume each statement to be true. Do not consider any relationships which might exist between the statements.

On the line following the statement place an X in the space after:

"qualified"--if the statement leads you to believe the author is qualified to write about the periodic Table.

"not qualified"--if the statement leads you to believe the author is not qualified to write about the Periodic Table.

In the space provided below each statement write a brief statement giving the reason(s) why you marked the answer you did.

86. He is a professor of Chemistry at an Ivy League
(Ev) University.
qualified ___ not qualified ___ no effect ___

87. He is not a member of the American Chemical Society.
(Ev) qualified ___ not qualified ___ no effect ___

88. He discovered the element Am.
(Ev) qualified ___ not qualified ___ no effect ___

89. He recently wrote a high school chemistry book.
(Ev) qualified ___ not qualified ___ no effect ___

90. He constructed a modern Periodic Table.
(Ev) qualified ___ not qualified ___ no effect ___

91. He is a teacher of high school chemistry and does not
(Ev) use a Periodic Table in his classes.
qualified ___ not qualified ___ no effect ___

92. He has a college degree in Russian History.
(Ev) qualified ___ not qualified ___ no effect ___

93. He wrote the article for an assignment in a college
(Ev) English class.
qualified ___ not qualified ___ no effect ___

94. He was educated in England.
(Ev) qualified ___ not qualified ___ no effect ___

95. He believes that only about half of the elements have
(Ev) been discovered.
qualified ___ not qualified ___ no effect ___

TABLE AS-1

ATOMIC STRUCTURE

Percentiles*

Raw Score	Knowledge 9	Knowledge 10	Knowledge 11	Knowledge 12	Comprehension 9	Comprehension 10	Comprehension 11	Comprehension 12
20	99	97	96	95				
19	95	91	87	85				
18	88	78	74	69				99
17	78	65	60	55		99	98	98
16	67	53	50	41		98	96	97
15	56	43	43	32		96	92	93
14	48	36	36	24	99	93	88	89
13	40	29	32	20	97	89	83	84
12	34	24	26	15	93	83	77	76
11	27	19	23	12	88	75	70	68
10	22	16	20	08	80	65	61	58
9	17	14	16	07	71	55	52	46
8	13	10	14	05	60	43	44	36
7	08	08	10		48	34	35	26
6	05	05	08	04	34	23	26	19
5	03	03	05	02	22	14	16	11
4	01	02	03	01	11	07	09	06
3		01	01		04	03	05	02
2						01	02	01
1					01		01	
0								

*Percentiles computed on the per cent of students scoring below a given raw score. Percentiles for this form were based on the following numbers of students:

Grade 9 - 1184
Grade 10 - 1419
Grade 11 - 1258
Grade 12 - 1196

TABLE AS-2

ATOMIC STRUCTURE

Percentiles

Raw Score	Application 9	Application 10	Application 11	Application 12	Analysis 9	Analysis 10	Analysis 11	Analysis 12
20								
19								
18			99	99				
17		99	98	98			99	
16		98	95	96			98	99
15		96	92	93		99	97	97
14	99	94	88	89		98	96	95
13	97	91	84	84		97	94	94
12	95	85	78	77	99	95	90	90
11	92	79	72	69	97	91	85	85
10	86	70	64	61	94	86	79	78
9	77	59	56	51	90	78	71	70
8	66	50	49	40	81	68	62	60
7	54	38	38	29	69	57	52	48
6	41	27	27	21	57	44	38	35
5	27	18	18	12	42	29	27	23
4	16	09	12	05	25	16	17	14
3	08	03	06	03	16	06	09	07
2	03	01	02	01	07	03	04	04
1	01		01		02	01	01	01
0								

TABLE AS-3

ATOMIC STRUCTURE

Percentiles

Raw Score	Synthesis 9	Synthesis 10	Synthesis 11	Synthesis 12	Evaluation 9	Evaluation 10	Evaluation 11	Evaluation 12
20								
19								
18								
17								
16								
15								
14								
13								
12								
11								
10								
9								
8						99		
7						98	99	99
6			99	99		96	97	96
5		99	98	98	99	93	95	91
4		97	96	95	97	87	91	87
3	99	94	91	89	92	81	83	74
2	96	88	84	79	82	77	72	64
1	89	77	72	63	68	61	59	49
0	71	59	50	40	48	38	41	30

TABLE AS-4

Item Data

Atomic Structure
Knowledge

Item No.	Key	Diffi-culty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
1	1	.88	.43	.43	.88	.04	07	.01	.00
2	1	.51	.43	.43	.51	.19	.09	.19	.02
3	2	.81	.46	.46	.11	.81	.06	.02	.00
4	3	.74	.54	.54	.08	.12	.74	.06	.01
5	2	.56	.36	.36	.19	.56	.10	.15	.01
6	2	.84	.58	.58	.06	.84	.05	.04	.00
7	4	.67	.56	.56	.05	.14	.21	.67	.00
8	3	.77	.54	.54	.05	.12	.77	.06	.00
9	4	.73	.59	.59	.10	[illegible]	.11	.73	.00
10	2	.58	.40	.40	.23	.58	.11	.07	.01
11	4	.78	.56	.56	.08	.05	.08	.78	.01
12	2	.73	.59	.59	.09	.73	.12	.05	.01
13	3	.77	.60	.60	.07	.06	.77	.10	.00
14	1	.85	.60	.59	.85	.05	.05	.03	.01
15	1	.87	.62	.62	.87	.04	.05	.03	.01
16	4	.25	.27	.27	.07	.13	.53	.25	.02
17	3	.87	.58	.57	.04	.05	.87	.03	.01
18	4	.63	.49	.49	.05	.09	.19	.63	.03
19	1	.60	.46	.46	.60	.12	.16	.11	.01
20	2	.62	.45	.44	.14	.62	.17	.05	.01

Raw Score

Mean = 14.05
S.D. = 4.25
KR #20 = 0.836

Corrected Score

Mean = 12.11
S.D. = 5.64
KR #20 = 0.929

N = 5105

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

TABLE AS-5

Item Data

Atomic Structure
Comprehension

Item No.	Key	Diffi-culty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
21	1	.52	.35	.40	.52	.11	.08	.27	.01
22	3	.53	.39	.42	.11	.10	.53	.24	.01
25	3	.62	.43	.48	.12	.17	.62	.08	.12
27	4	.52	.30	.34	.16	.15	.16	.52	.01
29	4	.67	.37	.43	.11	.10	.11	.67	.01
37	2	.67	.47	.50	.08	.67	.09	.13	.03
40	4	.09	.06	.04	.59	.13	.15	.09	.04
42	3	.18	.03	.00	.08	.21	.18	.47	.05
47	2	.46	.42	.40	.18	.46	.15	.13	.08
49	1	.59	.56	.55	.59	.10	.11	.12	.08
50	3	.56	.54	.52	.10	.14	.56	.11	.09
51	1	.35	.42	.40	.35	.16	.15	.24	.10
5[illegible]	4	.42	.50	.48	.12	.14	.20	.42	.11
58	2	.39	.50	.46	.13	.39	.23	.10	.15
65	3	.14	.03	.00	.18	.18	.14	.28	.22
70	2	.33	.42	.34	.13	.33	.18	.09	.27
72	4	.35	.49	.41	.11	.13	.13	.35	.28
73	3	.32	.44	.36	.12	.15	.32	.11	.30
75	4	.24	.39	.32	.14	.15	.16	.24	.31
79	1	.20	.31	.24	.20	.14	.18	.16	.32

Raw Score

Mean = 8.13
S.D. = 3.52
KR #20 = 0.693

Corrected Score

Mean = 5.00
S.D. = 4.38
KR #20 = 0.821

N = 5105

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

TABLE AS-6

Item Data

Atomic Structure
Application

Item No.	Key	Difficulty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
24	1	.53	.36	.39	.53	.16	.18	.12	.01
28	1	.67	.42	.44	.67	.18	.08	.06	.02
30	2	.31	.32	.33	.30	.31	.18	.19	.02
31	4	.51	.44	.49	.14	.16	.17	.51	.02
32	3	.53	.39	.41	.11	.11	.53	.22	.02
39	1	.45	.45	.47	.45	.13	.09	.28	.04
41	1	.40	.28	.27	.40	.41	.10	.05	.04
43	3	.63	.48	.47	.09	.14	.63	.08	.06
44	1	.35	.49	.49	.35	.26	.14	.19	.06
45	4	.31	.36	.36	.13	.34	.16	.31	.06
46	3	.31	.27	.24	.12	.18	.31	.30	.08
52	2	.22	.16	.12	.25	.22	.15	.27	.11
53	2	.50	.48	.43	.15	.50	.11	.14	.11
54	2	.38	.49	.46	.10	.38	.30	.11	.11
56	3	.36	.50	.47	.12	.29	.36	.11	.12
57	1	.25	.36	.32	.25	.14	.17	.30	.14
62	1	.33	.47	.43	.33	.21	.17	.11	.19
76	3	.28	.36	.30	.15	.17	.28	.08	.31
77	2	.25	.31	.25	.22	.25	.13	.08	.32
80	3	.22	.26	.20	.19	.12	.22	.15	.32

Raw Score

Mean = 7.79
S.D. = 3.60
KR #20 = 0.694

Corrected Score

Mean = 4.42
S.D. = 4.53
KR #20 = 0.826

N = 5105

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

TABLE AS-7

Item Data

Atomic Structure
Analysis

Item No.	Key	Diffi-culty	r_1 *	r_2 *	Proportion Choosing Each Response 1	2	3	4	Omit
23	1	.33	.25	.27	.33	.09	.17	.40	.01
26	3	.56	.31	.37	.11	.22	.56	.10	.01
33	2	.42	.39	.43	.20	.42	.16	.19	.03
34	4	.38	.43	.48	.18	.17	.23	.38	.04
35	1	.38	.37	.40	.38	.20	.22	.16	.04
36	1	.57	.32	.37	.57	.22	.08	.10	.03
38	4	.50	.39	.45	.12	.14	.20	.50	.04
48	2	.23	.30	.28	.23	.23	.22	.23	.09
59	2	.39	.42	.35	.08	.39	.21	.15	.16
60	3	.25	.39	.34	.26	.18	.25	.13	.17
61	1	.23	.33	.27	.23	.22	.24	.13	.18
63	4	.38	.50	.42	.12	.17	.13	.38	.20
64	2	.37	.54	.44	.11	.37	.17	.13	.21
66	1	.16	.24	.16	.16	.27	.20	.14	.23
67	4	.19	.30	.22	.16	.16	.25	.19	.24
68	4	.23	.41	.32	.17	.17	.16	.23	.26
69	2	.23	.35	.25	.19	.23	.20	.11	.27
71	4	.12	.19	.11	.17	.19	.24	.12	.28
74	1	.26	.39	.29	.25	.14	.14	.15	.31
78	3	.18	.19	.09	.20	.20	.18	.10	.32

Raw Scores

Mean = 6.35
S.D. = 3.17
KR #20 = 0.632

Corrected Score

Mean 2.83
S.D. 3.80
KR #20 = 0.760

N = 5105

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

APPENDIX B

Table of Contents

GLACIERS

Directions

This is a test of your ability to apply skills and understandings which you have been learning since you entered school. You will be asked to read a passage called "Glaciers" and then you will be asked to answer questions about it. You may refer to the passage at any time during the test.

Your answer must be marked on the answer sheet which has been provided for you. Use only the special pencil that has been provided. Do not make any stray marks. If you make an error, erase it completely before marking your new answer.

Make no marks in the test booklet.

The following is a sample question to show you how your answers are to be marked. Study the sample and if you have any questions raise your hand.

SAMPLE ON SEPARATE ANSWER SHEET

1. The article you will read is about
   1. glaciers.
   2. chemistry.
   3. glass-blowing.
   4. the Spanish-American War.

1. 1 ▬ 2 ___ 3 ___ 4 ___ 5 ___

You should answer as many questions as you can. Do not spend a great amount of time on any one question. If you are in doubt about the answer to a question, then guess. Remember that you may refer to the reading passage at any time during the test.

R.P. Kropp and H.W. Stoker, Editors
Department of Educational Research and Testing
and Institute of Human Learning
Florida State University
Tallahassee

April, 1965

Experimental edition; to be used for research purposes.

This experimental test was constructed as part of a research project supported by the U. S. Office of Education under Contract Number OE-4-10-019, Project Number 2117.

GLACIERS

Ice plays a critical role in the water economy of the earth. About 86 per cent of it is in the Antarctic, where it exerts a profound influence on the weather in all parts of the world.

By William O. Field

Water is one of the few substances on earth existing in nature in all three physical states--liquid, solid, and gaseous. Altogether our planet contains some 350 million cubic miles of water, most of it, of course, in the oceans. Of the earth's total water budget, not much more than one per cent is in the solid form of ice or snow, and far less than that in the form of water vapor in the atmosphere. Yet these proportions make up a delicate balance which is immensely important to life on the earth. Any appreciable change in the ratios of water, ice and atmospheric moisture would have catastrophic consequences for man and his economy. The ice piled in glaciers on the lands, for instance, exercises a vital control over sea levels, climate and the continents' water supplies.

Glaciers now cover about ten per cent (nearly six million square miles) of the world's land area. Our estimate of the total amount of water in them is only a rough guess, mainly because we have only a hazy notion of the thickness of the Antarctic ice sheet. This vast icecap accounts for about 86 per cent of the world's glacial area. The Greenland icecap makes up another ten per cent. The remaining four per cent is not minor, as far as its effects go, for it includes tens of thousands of square miles of glaciers on mountains in the temperate zones, where they intimately influence man's climate and water supplies.

Estimates of the total volume of water in the world's glaciers range from about 2.4 million to more than six million cubic miles. If all this ice melted, the level of the world's oceans would rise by something like 65 to 200 feet!

Glaciers can grow only in areas where the snowfall is great enough year after year to exceed the annual rate of melting. Consequently, the ice sheet is not necessarily thickest where the climate is coldest. In Alaska the greatest concentration of glaciers is along the southern coast, which is the warmest part of the Territory but has the heaviest winter snowfall. Parts of northern Greenland are barren of glaciers because there is not enough snowfall.

As snow accumulates, the pressure of the mountainous layers compacts it into ice. Under its own weight ice begins to flow to lower elevations. The rate of flow of glaciers varies tremendously: some move very slowly while others slide as much as 50 feet per day during the summer. At the lower elevations, the glacier melts or discharges icebergs into the sea. But under suitable conditions, the glacier front may advance over the land year after year. It takes only a slight change in the combination of annual snowfall, melting-season temperatures and other meterological conditions to produce an advance or retreat of a glacier.

Probably during most of the earth's history it has been free of glaciers. We are in an exceptional era--neither glacial nor nonglacial. During the last million years there have been at least four great ice ages; at their maximum, ice covered about 32 per cent of the world's land surface. The ice ages were separated by long warm intervals during which the glaciers nearly disappeared. At present we seem to be in an in-between stage, somewhere between a glacial and an interglacial age. Some glaciers are growing; others are disappearing.

During the last Ice Age the sea level probably was more than 300 feet lower than now. Over the world the temperatures averaged 7 to 14 degrees colder. There were five continental ice sheets of more than one million square miles each. Three of these, in North America, Europe and Siberia, have disappeared, but the two in Greenland and Antarctica remain. Mountain glaciers have all shrunk.

Human civilizations began to arise in Western Asia and North Africa just as the European and North American sheets were disappearing. About 3000 B.C. the climate in many, if not all, parts of the world was drier and warmer by two or three degrees than at present. The sea level was apparently five to six feet higher. The glacial region in the Alps was at least 1,000 feet higher than today. Ice in the Artic Ocean probably melted completely each summer. Parts of the temperature regions where small mountain glaciers now furnish the summer water supply must have been arid.

Conditions began to change drastically about 1000 B.C. The climate became colder and more stormy in many parts of the world, and by about 500 B.C. glaciers began to grow again. Then, in the first millennium of the Christian era, came a period of glacier recession. After that glaciers advanced again to a maximum in the 17th to 19th centuries. This resurgence of glaciers was noted directly by observers

in the Alps, Scandinavia and Iceland. Since the latter half of the 19th century, glaciers throughout the world have tended to shrink once again. As a result the sea level has apparently been rising recently at the rate of approximately 2.5 inches per century. Some glaciers, however, have advanced, contrary to the general trend. In parts of the western U.S. there is a growth of glaciers at present which may indicate a changing climate.

Glaciers have been studied seriously for a little more than 100 years. Beginning in 1919 Hans Wison Ahlmann of the University of Stockholm (now Sweden's Ambassador to Norway) introduced a new era in glaciology. He took a new look, in greater detail, at glaciers in Scandinavia, Iceland, Spitsbergen and northeast Greenland, and his examination led to new methods of measuring their nourishment and wastage. Observations of glaciers are now being made on a systematic basis in several parts of the world. During the last decade, important studies have been carried out in Greenland, especially by Paul Victor's French Polar Expeditions, which determined the volume of the Greenland ice sheet and studied its regimen over a broad area.

The little-known Antarctic ice sheet is more than one and a third times the size of the U.S. and its territories. It covers practically the whole continent of Antartica. Fully three million square miles of the continent have never been seen even from the air. The continent's icecap is known to rise as high as 10,000 feet, but the thickness of the ice has been measured in only a few places.

Table 1. Distribution of Water Volume

LOCATION	CUBIC MILES
Water in the oceans (close estimate)	329,000,000
Water in the atmosphere (rough estimate)	3,600
Water in glaciers (average of high and low estimates)	4,200,000
Water in lakes and rivers (rough estimate)	55,000
Ground water above 12,500 feet (very rough estimate)	1,080,000
Ground water below 12,500 feet (very rough estimate)	19,700,000

Table 2. Distribution of Ice by Area

LOCATION	SQUARE MILES
Africa	8
Antarctica	5,019,000
Asia	42,200
Canadian Arctic Islands	45,000
Europe	4,370
Greenland	666,300
North America	30,890
Northern Atlantic and European Arctic Islands	45,400
Pacific Islands	392
South America	9,650
Sub-Antarctica Islands	1,160
World Total	5,864,370

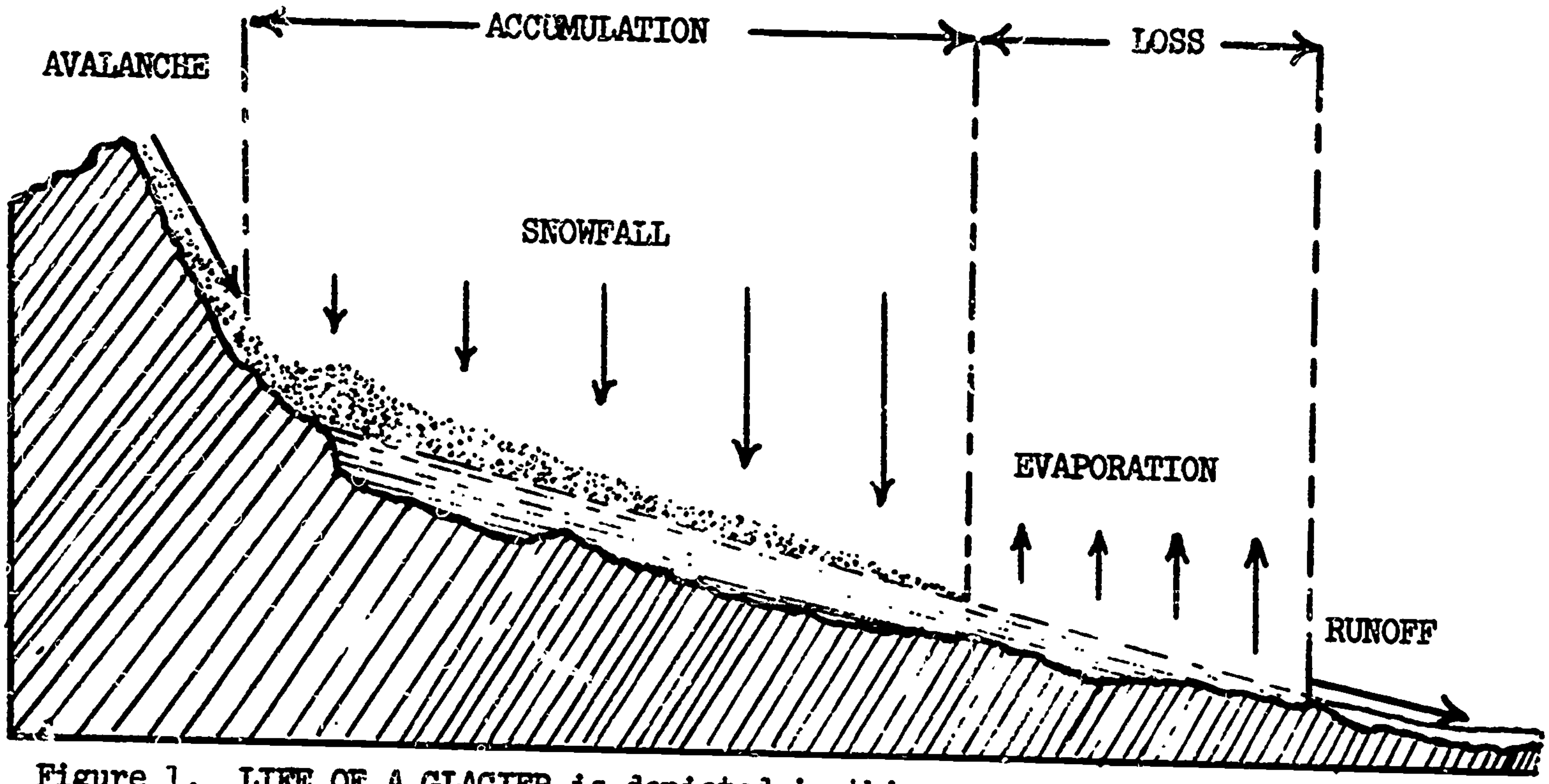


Figure 1. LIFE OF A GLACIER is depicted in this cross section of an ideal valley glacier. Falling snow carried by avalanche is compressed into ice, which begin to move by its own weight. The line dividing the areas of accumulation and loss is the firn line, where total accumulation equals total melting. Variations in snowfall, temperature and other conditions determine whether the glacier advances.

Adapted from

William O. Field, "Glaciers," Scientific American, September, 1955, pp. 84-92.

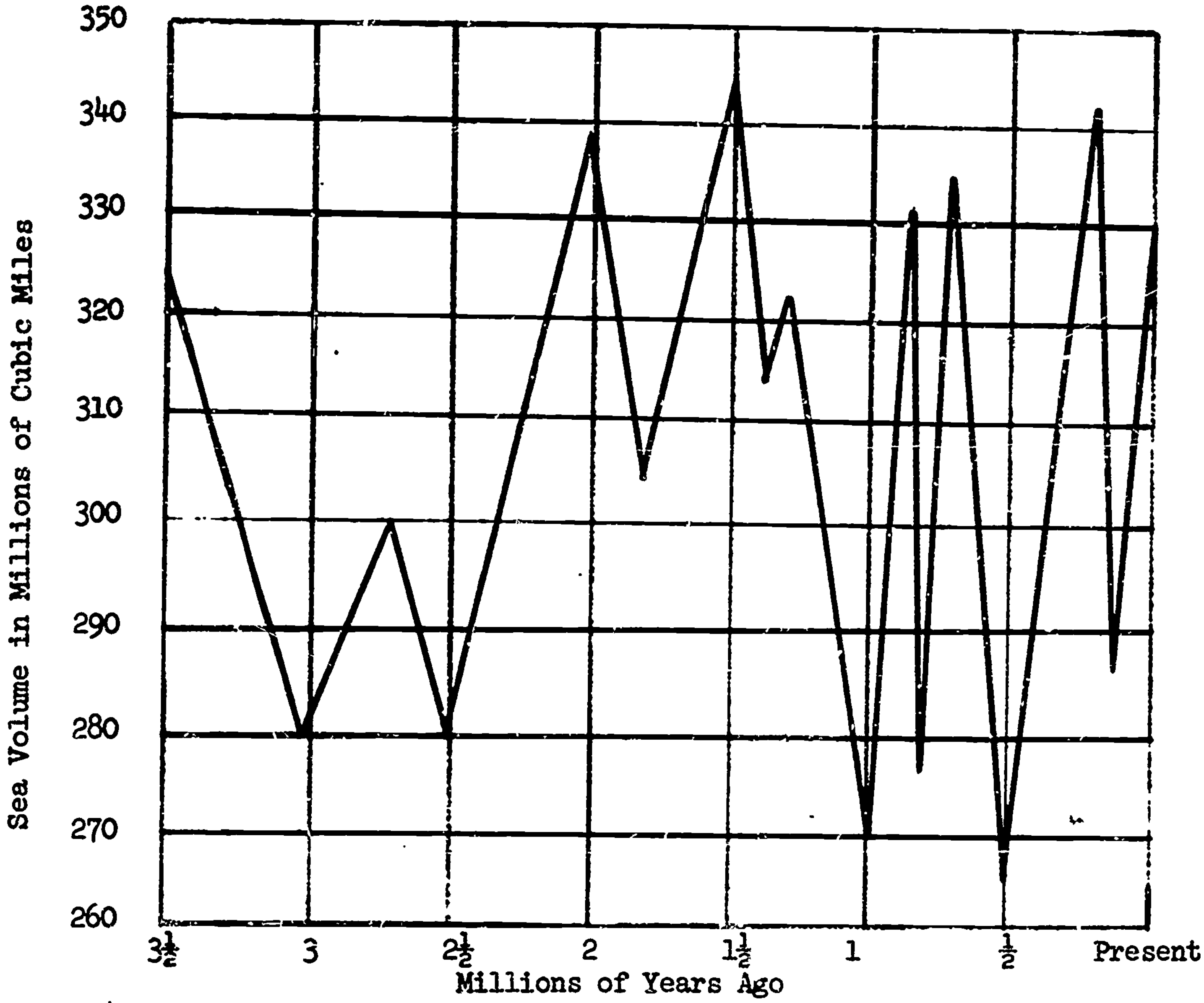


Figure 2. Hypothetical Variations in Sea Volume

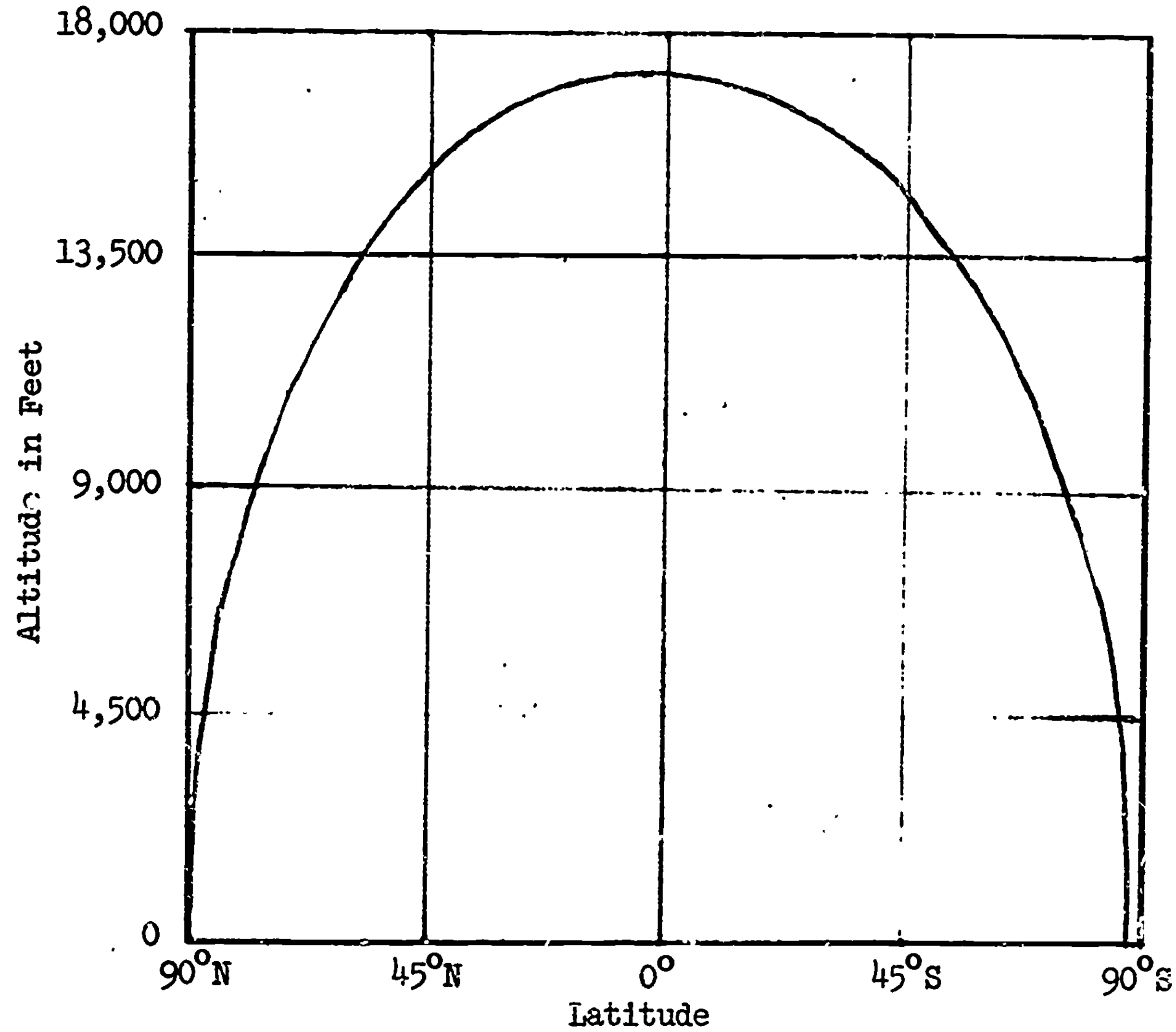


Figure 3. Approximate Height of Snow Line at Various Latitudes

1. The glaciers in the Antartic influence the weather
(K) 1. only in the southern hemisphere.
2. only in the northern hemisphere.
3. primarily in the northern hemisphere.
4. in all parts of the world.

2. Which of the following is one of the natural physical states of matter?
(K) 1. ice
2. frozen
3. solid
4. hard

3. In which physical state is most of the earth's water?
(K) 1. liquid
2. frozen
3. gaseous
4. ice

4. What per cent of the earth's water budget exists as water vapor?
(K) 1. less than 1
2. 4
3. 10
4. 86

5. The ice in glaciers does <u>not</u> control
(K) 1. sea level
2. climate
3. wind speed
4. water supply

6. Glaciers now cover
(K) 1. about 10 per cent of the world's land area.
2. about 10 per cent of the world's surface area.
3. about 60 million square miles.
4. only the polar regions.

7. The world's second largest glacial area is in
(K) 1. Antarctica
2. Canada
3. Greenland
4. Siberia

8. The thickness of glaciers is primarily determined by the
(K) 1. land surface.
2. altitude of the glacier.
3. average temperature of the area.
4. amount of snowfall.

9. There are fewer glaciers in northern Alaska than in southern Alaska because northern Alaska has
(K) 1. colder weather.
2. less snowfall.
3. more snowfall
4. fewer valleys in which glaciers could form.

10. Icebergs are formed by
(K) 1. the freezing of sea water.
2. glaciers advancing into the oceans.
3. glaciers forming over the oceans.
4. snow falling into the oceans.

11. How many ice ages have occurred during the last million years?
(K) 1. 1
2. 2
3. 3
4. 4

12. Approximately what per cent of the world's land surface was covered with ice during the great ice ages?
(K) 1. 10
2. 20
3. 30
4. 40

13. We now live in a period which is neither glacial nor nonglacial. Which of the following is the best evidence of this fact?
(K) 1. Extremely large glaciers are melting.
2. The climate is constantly changing.
3. The annual snowfall is decreasing.
4. Some glaciers are growing and others are shrinking.

14. Of the five continental ice sheets formed during the last ice age two remain. One of them is located in
(K) 1. Canada
2. Greenland
3. Norway
4. Siberia

15. As compared with our present climate, the climate of the world in 3000 B.C. was
(K) 1. drier and colder.
2. drier and warmer.
3. damper and colder.
4. damper and warmer.

16. Hans Ahlmann did <u>not</u> study the glaciers in
(K)
1. Antarctica
2. Greenland
3. Iceland
4. Scandinavia

17. Where is most of the world's water?
(K)
1. in the oceans
2. in the atmosphere
3. below the ground
4. in glaciers

18. According to Table 2, glaciers can be found in
(K)
1. Africa
2. Europe
3. South America
4. all of these places

19. The line dividing the areas of accumulation and loss is called the
(K)
1. firn line
2. dew line
3. glacier line
4. loss-runoff line

20. According to Figure 2, what was the maximum volume of water and when did it occur?
(K)
1. 265 million cubic miles, $\frac{1}{2}$ million years ago.
2. 345 million cubic miles, $1\frac{1}{2}$ million years ago.
3. 339 million cubic miles, 2 million years ago.
4. 270 million cubic miles, 1 million years ago.

21. If glaciers cover about ten per cent of the land area, or about six million square miles of the earth, then the land area of the earth, in square miles, is approximately
(C)
1. 5,400,000
2. 6,000,000
3. 54,000,000
4. 60,000,000

22. The great continental ice sheets disappeared from the United States and Europe about
(C) 1. 3,000 years ago
2. 5,000 years ago
3. 250,000 years ago
4. 500,000 years ago

23. If the average temperature in the Great Lakes region dropped until the lakes were never completely free of ice, which of the following would most likely result?
(An) 1. Valley glaciers would form in the lake beds.
2. The amount of water in the oceans would increase.
3. The amount of snow in the southern United States would increase
4. The Artic ice sheet would recede.

24. What might happen if the earth's glaciers were to melt during the next decade?
(Ap) 1. Greenland would develop into an important nation.
2. The great seaports of the world would disappear under water.
3. Antarctica would emerge as a large continental mass.
4. Australia would submerge.

25. The term "RUNOFF" which appears in Figure 1
(C) 1. is defined clearly in the reading passage.
2. is not essential for understanding glaciers.
3. restates the central idea of glacier formation.
4. is a synonym for "melt."

26. If one wished to estimate from the information in the passage the amount of water in the Antarctic ice sheet, then what assumption would have to be made?
(An) 1. The average thickness of the Antarctic ice sheet is equal to the average thickness of all the world's glaciers.
2. The average temperature of Antarctica is about the same as the average temperature of other glacial areas.
3. Snowfall in Antarctica is about equal to the snowfall in other areas.
4. The ratio of growth of the Antarctic ice sheet is about the same as that of other glaciers.

27. The fact that the great Alaskan glaciers are located in forested areas implies that these glaciers are
(C) 1. part of a vast ice cap.
2. the world's thickest glaciers.
3. remaining portions of a receding ice cap.
4. temperate zone glaciers.

28. What is the relationship between
(Ap) A. Altitude of snow line, and
B. Average rainfall.
1. An increase in A is accompanied by an increase in B.
2. An increase in A is accompanied by a decrease in B.
3. A and B are unrelated.
4. The relationship cannot be determined from the passage.

29. During the last ice age, the highest summer temperatures in what is now the southeastern United States was approximately
(C) 1. 70 degrees
2. 85 degrees
3. 110 degrees
4. 120 degrees

30. What is the relationship between
(Ap) A. Distance from the Poles, and
B. Average wind velocity.

1. An increase in A is accompanied by an increase in B.
2. An increase in A is accompanied by a decrease in B.
3. A and B are unrelated.
4. The relationship cannot be determined from the passage.

31. Reread the third paragraph of the article. Now assume that 5 million cubic miles of water are contained in the world's glaciers. If the glaciers melted completely, approximately how many feet would ocean levels rise.
(Ap) 1. 70
2. 100
3. 160
4. 190

32. If the glacier shown in Figure 1 is located at latitude 68 degrees N, according to the information given in Figure 3, "RUNOFF" would begin at approximately
(Ap) 1. 4,500 feet
2. 9,000 feet
3. 10,500 feet
4. 13,500 feet

33. Which of the following best describes the entire article?
(An) 1. It presents evidence that we are in an interglacial period.
2. It describes in scientific language the growth and decline of glaciers.
3. It presents arguments to show the need for increased study of glaciers.
4 It shows the critical role ice plays in the economy of the earth.

34. Suppose that one inch of snow per year fell in a barren land, and that the temperature never exceeded the freezing point, and that no glaciers ever formed. One possible explanation for the absence of glaciers is
(An) 1. the absence of avalanches.
2. that the snow evaporates.
3. that runoff is greater than snowfall.
4. that the snow is immediately converted to ice.

35. The <u>least</u> important factor in glacier formation is
(An) 1. annual snowfall.
2. low temperature.
3. atmospheric pressure.
4. distance from the equator.

36. If you saw a piece of ice, weighing about 100 pounds, floating in the Gulf of Mexico, which of the following would be the most reasonable explanation for its presence?
(An) 1. It was artifically formed and had just been dropped from a passing boat.
2. It was an iceberg which was formed off the coast of Greenland and drifted to the Gulf.
3. It was formed in the Rocky Mountains and was carried to the Gulf.
4. It was the result of hail massing as it fell into the Gulf.

37\. The size and number of glaciers have remained relatively unchanged during the past two thousand years; however,
(C) 1. there have been substantial changes from century to century during the period.
2. sea level has increased by $2\frac{1}{2}$ inches each century for the last five centuries.
3. glaciers in the United States have disappeared but those in Europe have increased in number and size.
4. the volume of water in the atmosphere has increased substantially.

38\. Temperate zone glaciers play an important role in the study of climatology because
(An) 1. they are easily accessible for observation and study.
2. there are plentiful records of weather conditions and glacial growth for this zone.
3. the physical characteristics of temperate zone glaciers are different from those of glaciers of other zones.
4. the temperate zone glaciers are more likely to advance or retreat.

39\. The process of converting snow to ice in the formation of a glacier is similar to the process of making
(Ap) 1. steam from water.
2. wool by sheep.
3. rayon from chemicals.
4. coal from vegetation.

40\. According to Figure 2, the smallest amount of water in solid form existed how many million years ago?
(C) 1. $\frac{1}{2}$
2. 1
3. $1\frac{1}{2}$
4. 2

41\. What is the relationship between
(Ap) A. Amount of glaciation in the world, and
B. Surface area of land.
1. An increase in A is accompanied by an increase in B.
2. An increase in A is accompanied by a decrease in B.
3. A and B are unrelated.
4. The relationship cannot be determined from the passage.

42. If all of the ice in Antarctica were melted by artificial means and all other ice areas remained the same then the maximum rise in the world's oceans would be
(C) 1. 60 feet
2. 170 feet
3. 240 feet
4. 210 feet

* * * * * * *

NOTE: The following information applies to questions 43-45. Assume the following about the conditions of a geographic area:
A. average annual snowfall is 10,000 cubic feet per square mile.
B. average daily high temperature in 70 degrees F. in the summer months.
C. average annual melt is 10,000 cubic feet per square mile.
D. these conditions have existed for 100 years.

43. This region probably
(Ap) 1. has no glaciers.
2. is highly glaciated.
3. is near a glaciated area.
4. has a few small glaciers.

44. If the annual melt were to be reduced to 9,500 cubic feet per year, how would the conditions change?
(Ap) 1. The annual snowfall would decrease to less than 9,500 cubic feet per square mile.
2. Glaciers would begin to form.
3. The annual snowfall would increase.
4. The glaciers would begin to disappear.

45. What would cause a decrease in the annual melt?
(Ap) 1. Increase in annual snowfall.
2. Decrease in average daily summer temperature.
3. Decrease in average atmospheric pressure.
4. Increase in average daily winter temperature.

* * * * * * *

46. What is the relationship between
(Ap) A. Glaciation of North America, and
B. Prevailing temperature in southeast United States.
1. An increase in A is accompanied by an increase in B.
2. An increase in A is accompanied by a decrease in B.
3. A and B are unrelated.
4. The relationship cannot be determined from the passage.

47. The purpose of Figure 1 is to
(C) 1. present a picture of glacial formation.
2. show areas of a glacier where snowfall and evaporation occur.
3. introduce the concept of runoff.
4. show one of the results of an avalanche.

48. Reread the third paragraph of the passage. Which of the following facts must be known to justify the conclusion?
(An) 1. The average world temperature.
2. The volume of water in the oceans.
3. The volume of ice in glaciers.
4. The average snowfall per year.

49. According to Figure 2, which of the following best describe the world 1½ million years ago.
(C) 1. It was almost completely covered with glaciers.
2. There were deep valley glaciers in high mountainous regions.
3. There were probably no glaciers at all.
4. Continental ice sheets were growing.

50. The Antarctic ice sheet is approximately the size of
(C) 1. Europe and Asia.
2. the United States and Russia.
3. the United States and Mexico.
4. the Arctic ice sheet.

51. Which of the following best describes conditions during the past 2,000 years?
(C) 1. Glaciation has fluctuated.
2. Sea level has been constantly decreasing.
3. America's continental glaciers have disappeared, but Europe's have not.
4. Water in the atmosphere has decreased.

* * * * * * *

NOTE: The following information must be used in answering questions 52-54:

The earth has existed many millions of years during which its ocean content was not always 330 million cubic miles. From the preceding statement and what you know of earth history, answer questions 52-54 according to the following key:

KEY:
1. Water content exceeded 330 million cubic miles.
2. Water content was less than 330 million cubic miles.
3. Water content remained 330 million cubic miles.
4. No evidence exists for a valid conclusion.

52. During the time of the last ice sheet in North America. (Ap)

53. During the time when most of the Caribbean Islands were larger than they are now. (Ap)

54. When the volume of water in the Great Lakes was at its maximum. (Ap)

* * * * * * *

55. A glacier can best be described as a (C)
    1. snow field.
    2. moving ice field.
    3. product of low temperature.
    4. valley packed with ice.

56. What is the relationship between (Ap)
    A. Average world temperature, and
    B. Amount of water in the oceans.
    1. An increase in A is accompanied by an increase in B.
    2. An increase in A is accompanied by a decrease in B.
    3. A and B are unrelated.
    4. The relationship cannot be determined from the passage.

57. What is the relationship between
(Ap) A. Amount of snowfall, and
B. Amount of glaciation.
1. An increase in A is accompanied by an increase in B.
2. An increase in A is accompanied by a decrease in B.
3. A and B are unrelated.
4. The relationship cannot be determined from the passage.

58. Approximately 96 per cent of the glacial area of the world is located
(C) 1. in the southern hemisphere.
2. outside the temperate zone.
3. in the eastern hemisphere.
4. in Antarctica.

59. One plausible reason for the lack of glaciers in Africa is that
(An) 1. it has no high mountains.
2. the southern hemisphere is less glaciated than the northern hemisphere.
3. large areas of desert prevent glaciers from forming.
4. the average annual range of temperature is too small.

60. An increase of 10 per cent in the average world temperature would have which of the following effects?
(An) 1. Increase the water in the atmosphere.
2. Cause over 4,500,000 cubic miles of water to freeze into glaciers.
3. Lessen the 300 million cubic miles of water in the oceans.
4. Decrease average annual rain fall.

61. Water at the northern end of the Greenland ice sheet would most probably exist naturally in
(An) 1. solid form only.
2. solid and liquid forms only.
3. solid and gaseous forms only.
4. solid, liquid and gaseous forms.

62. What is the relationship between
(Ap) A. Number of glaciers near the coast, and
B. Number of icebergs in the ocean.
1. An increase in A is accompanied by an increase in B.
2. An increase in A is accompanied by a decrease in B.
3. A and B are unrelated.
4. The relationship cannot be determined from the passage.

63. Assume that the average temperature in the United States will drop 50 degrees F. tomorrow, and remain the same for 500 years. Where would you expect to see the first new valley glaciers?
(An) 1. Kansas-Nebraska wheat belt.
2. Texas-Oklahoma cattle country.
3. Northern Alaska.
4. Central eastern seaboard.

64. If all glaciers melted, which one of the following predictions would be false?
(An) 1. Coastal resorts would be flooded.
2. The climate would be warmer.
3. The salt content of the oceans would be increased.
4. Mass migration of people would take place.

65. The total volume of water in the world's glaciers is approximately
(C) 1. one million cubic miles.
2. two million cubic miles.
3. four million cubic miles.
4. seven million cubic miles.

66. Suppose the Anarctica ice sheet were melted. What would be the new elevation of a city which is now 160 feet above sea level?
(An) 1. 180 feet below sea level
2. 50 feet below sea level
3. 25 feet above sea level
4. 50 feet above sea level

67. The life of a glacier is depicted in Figure 1. None of the ice mass would occur in the areas marked
(An) 1. avalanche and runoff.
2. accumulation and avalanche.
3. accumulation and runoff.
4. loss and runoff.

68. Table 1 shows the distribution of water on the earth. Which of the following sources of water is not accounted for in this table?
(An)
1. Water "trapped" in coal and oil deposits.
2. Water contained in living plants and animals.
3. Water contained in clouds.
4. Water contained in preci itation.

69. Figure 2 indicates that the change in volume of water has been constant during the past two hundred thousand years. However, this figure seems to conflict with the discussion concerning glacier activity during the past three thousand years. What might account for this disagreement?
(An)
1. The author, W. O. Field, is probably wrong in his description of glacier activity.
2. The graph (Figure 2) is probably wrong.
3. The changes in water volume were not affected by the glacier changes during the past 3,000 years.
4. The fluctuations in water volume during the past 3,000 years were insignificantly small in comparison with the long range trend.

70. The serious study of glaciers began about the time of the
(C)
1. Civil War.
2. Golden Age of Greece.
3. French Revolution.
4. discovery of America.

71. What is the primary reason that glaciers do not exist in central Canada?
(An)
1. The altitude of this region is less than 13,500 feet.
2. The temperature reaches 80 degrees F. in the summer months.
3. The temperature reaches 30 degrees F. below zero.
4. It is too far from the polar ice cap.

72. Glaciers stopped advancing over North America when
(C)
1. they reached the mild climate south of the Tropic of Cancer.
2. accumulation exceeded loss.
3. the snowfall failed to keep pace with the annual melt.
4. the edges reached the sea and formed icebergs.

73. According to Figure 2, how did world conditions one million years ago compare to those of $1\frac{1}{2}$ million years ago?
(C)
1. Similar, cold and dry.
2. Similar, warm and humid.
3. Dissimilar, almost opposite.
4. Not enough information is given to make comparison.

74. Suppose that a scientist calculated that if a certain proportion of the glaciers melted, then a city now 500 feet above sea level would be only 250 feet above sea level. He probably assumed which one of the following in making his calculations?
(An)
1. The volume of a kilogram of water varies directly with temperature.
2. One cubic foot of ice, when melted, produces one cubic foot of water.
3. Some of the melted ice would remain in mountain lakes.
4. The slope of shore lines is relatively constant around the world.

75. From the information given in Figure 3, glaciers could form at 45 degrees north latitude provided that
(C)
1. there is enough pressure.
2. the altitude is 15,000 feet or higher.
3. there are no forested areas.
4. the temperature never gets low enough to prevent snowfall.

76. What is the relationship between
(Ap)
A. Melting of the Antarctic ice sheet, and
B. Area of North Dakota.
1. An increase in A is accompanied by an increase in B.
2. An increase in A is accompanied by a decrease in B.
3. A and B are unrelated.
4. The relationship cannot be determined from the passage.

77. Refer to Table 1. The ratio of water in glaciers to water in the oceans is about
(Ap)
1. 1/75.
2. 1/15.
3. 1/8.
4. 1/30.

78. Suppose that conditions were changed so that the volume of water in one of the categories in Table 1 was decreased by 500 cubic miles. In which one of the following categories would this decrease result in the most drastic climatic changes?
(An) 1. water in oceans
2. water in atmosphere
3. water in glaciers
4. ground water above 12,500 feet

79. Which change would make Figure 1 more descriptive of valley glacier formation?
(C) 1. Add snowfall arrows to include the entire figure.
2. Extend accumulation line to include the area marked evaporation.
3. Add snowfall arrows to include the area marked avalanche.
4. Add evaporation arrows to include the area marked runoff.

80. The author states that we are living in an exceptional era with regard to glaciation. Which one of the following would substantiate this statement the least from the physical standpoint?
(Ap) 1. A new continental glacier is forming.
2. The world is partially glaciated.
3. The Arctic Ocean is open water.
4. A systematic study of glaciers is taking place.

GLACIERS
(Booklet Number 2)

Directions

Your answers to each of items 81-95 should be written directly on the test booklet in the space provided below each item. Should you need additional space for an item write "continued" then write the remainder of your answer on the back of the last page of the booklet.

You received with this test booklet another booklet which contains the article about Glaciers. You may refer to it any time during the examination. You might wish to scan the article after you read these directions to refresh your memory.

Your answers to these items will depend on your understanding and thoughts about the article. Each of you might write an entirely different answer to an item because there are many possible answers to each of them. So please write whatever you believe to be appropriate.

Answers to these items do not appear as such in the article. Therefore, you will certainly answer incorrectly if you copy material directly from the article.

81. (Sy) Suppose that a glacier is advancing toward a town and that you have been asked to estimate how long it will take for it to reach the town. Briefly describe the steps you would take to make this estimate.

82. (Sy) A scientist believes that advancing glaciers slide on a thin layer of water which lies between the bottom of the glacier and the earth. Outline a plan to test whether his belief is true or false.

83. (Sy) The land over which a glacier has passed suffers considerable damage. Briefly outline why glaciers cause damage and describe the kind of damage they might cause.

84. (Sy) Briefly describe a method by which you could determine if the vertical thickness of a glacier is increasing or decreasing. In your method, indicate measurements you would take and how would you use them: Remember that glaciers move!

85. (Sy) Suppose you knew that an advancing glacier would reach town in 12 months. Outline a plan which could be followed to prevevent the glacier from reaching the town.

86.-95. Suppose that you have been asked to evaluate the author's qualifications to write an article on Glaciers. Listed below are several statements which could be true about the author. Assume each statement to be true. Do not consider any relationships which might exist between the statements.

On the line following the statement place an X in the space after:

"qualified"--if the statement leads you to believe the author is qualified to write about glaciers.

"not qualified"--if the statement leads you to believe the author is not qualified to write about glaciers.

"no effect"--if you believe the statement write a brief statement giving the reason(s) why you marked the answer you did.

86. He is a professor of English at a large university.
(Ev) qualified ___ not qualified ___ no effect ___

87. His books are often quoted in newspapers and magazines.
(Ev) qualified ___ not qualified ___ no effect ___

88. He spent a year at Thule Air Force Base in Greenland.
(Ev) qualified ___ not qualified ___ no effect ___

89. This was the first article written by the author
(Ev) published in Scientific American.

90. His books are often quoted by other glaciologists.
(Ev) qualified ___ not qualified ___ no effect ___

91. He won a Nobel prize for his work on the geology of
(Ev) Haiti.
qualified ___ not qualified ___ no effect ___

92. He is a professor of meteorology (the study of weather)
(Ev) at a large university.
qualified ___ not qualified ___ no effect ___

93. His name is not included in the Directory of American
(Ev) Scholars.
qualified ___ not qualified ___ no effect ___

94. He is an amateur mountain climber.
(Ev) qualified ___ not qualified ___ no effect ___

95. He is a recognized expert in the field of electronics.
(Ev) qualified ___ not qualified ___ no effect ___

TABLE G-1

GLACIERS

Percentiles*

Raw Score	Knowledge 9	Knowledge 10	Knowledge 11	Knowledge 12	Comprehension 9	Comprehension 10	Comprehension 11	Comprehension 12
20	88	83	76	76				
19	74	64	57	53				
18	60	50	42	38			99	99
17	48	38	31	27		99	97	97
16	39	29	23	19	99	97	95	93
15	31	23	17	14	97	94	90	88
14	25	17	13	10	96	90	85	82
13	19	14	09	07	93	84	78	76
12	15	11	07	05	87	77	70	65
11	12	08	05	04	81	69	61	56
10	09	06	04	03	71	59	49	47
9	06	05	03	02	61	49	37	37
8	04	04			49	37	29	27
7	02	03	02		37	27	20	19
6		02		01	26	18	12	11
5	01	01	01		16	10	07	06
4					10	04	04	03
3					04	02	01	01
2					02	01		
1					01			

*Percentiles computed on the per cent of students scoring below a given raw score. Percentiles for this form were based on the following numbers of students:

Grade 9 - 1223
Grade 10 - 1431
Grade 11 - 1303
Grade 12 - 1186

TABLE G-2

GLACIERS

Percentiles

Raw Score	Application 9	Application 10	Application 11	Application 12	Analysis 9	Analysis 10	Analysis 11	Analysis 12
20								
19				99				
18	99	99	98	97				
17	98	96	94	94				
16	96	93	90	89				
15	94	88	82	82			99	99
14	90	83	74	74		98	98	98
13	84	76	65	65	99	96	95	95
12	78	68	56	56	97	92	90	89
11	69	60	47	48	93	86	82	81
10	62	51	39	40	88	76	72	68
9	53	42	32	30	78	67	60	56
8	45	33	25	24	67	54	49	43
7	33	24	17	17	55	41	36	32
6	23	16	11	11	41	29	25	24
5	14	09	06	06	27	18	16	15
4	08	04	04	03	18	11	09	07
3	04	02	02	01	09	05	04	03
2	01	01	01		03	02	02	01
1					01			

TABLE G-3

GLACIERS

Percentiles

Raw Score	Synthesis 9	Synthesis 10	Synthesis 11	Synthesis 12	Evaluation 9	Evaluation 10	Evaluation 11	Evaluation 12
20								
19								
18								
17								
16								
15								
14								
13								
12							99	99
11						99	98	97
10		99	99	99	99	97	95	96
9	99	98	98	98	97	94	92	91
8	98	97	96	96	93	89	84	82
7	97	93	91	91	87	81	74	72
6	94	90	86	86	78	72	65	60
5	89	81	77	77	67	62	54	46
4	79	69	63	63	55	49	42	36
3	65	54	48	47	45	40	33	28
2	46	35	32	29	34	33	26	21
1	25	19	17	15	25	26	19	16
0	10	08	06	05	17	21	14	12

PRECEDING PAGE BLANK- NOT FILMED

TABLE G-4

Item Data

Glaciers
Knowledge

Item No.	Key	Difficulty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
1	4	.78	.43	.43	.08	.03	.10	.78	.01
2	3	.81	.39	.39	.13	.04	.81	.02	.00
3	1	.90	.36	.35	.90	.03	.03	.04	.00
4	1	.80	.54	.54	.80	.10	.06	.04	.00
5	3	.89	.46	.46	.03	.05	.89	.03	.00
6	1	.83	.52	.52	.83	.09	.05	.03	.00
7	3	.84	.45	.45	.12	.02	.84	.02	.00
8	4	.79	.41	.41	.04	.10	.06	.79	.01
9	2	.74	.43	.43	.06	.74	.14	.06	.00
10	2	.80	.44	.44	.06	.80	.06	.08	.01
11	4	.79	.57	.57	.04	.06	.11	.79	.00
12	3	.84	.57	.57	.05	.04	.84	.07	.01
13	4	.86	.48	.48	.03	.07	.03	.86	.00
14	2	.89	.46	.46	.02	.89	.02	.07	.00
15	2	.77	.52	.52	.09	.77	.12	.02	.00
16	1	.85	.57	.57	.85	.03	.05	.07	.00
17	1	.91	.45	.45	.91	.02	.01	.06	.00
18	4	.89	.44	.43	.03	.05	.03	.89	.00
19	1	.75	.56	.56	.75	.04	.10	.11	.00
20	2	.86	.47	.47	.05	.86	.06	.03	.00

Raw Scores

Mean = 16.593
S.D. = 3.517
KR #20 = 0.816

Corrected Scores

Mean = 15.473
S.D. = 4.676
KR #20 = 0.919

N = 5210

*r_1 = Point-biserial correlation, raw scores
r_2 = Point-biserial correlation, corrected scores

TABLE G-5

Item Data

Glaciers
Comprehension

Item No.	Key	Difficulty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
21	4	.67	.40	.41	.05	.12	.16	.67	.01
22	2	.30	.30	.32	.49	.30	.12	.07	.02
25	4	.73	.24	.25	.07	.11	.08	.73	.01
27	4	.42	.27	.28	.14	.07	.37	.42	.01
29	2	.35	.17	.18	.49	.35	.12	.04	.01
37	1	.33	.20	.20	.33	.44	.12	.10	.01
40	3	.16	.26	.26	.70	.08	.16	.05	.01
42	2	.40	.40	.42	.32	.40	.14	.12	.01
47	1	.63	.42	.43	.63	.16	.10	.09	.02
49	3	.30	.40	.41	.40	.11	.30	.16	.02
50	3	.56	.44	.44	.13	.13	.56	.14	.03
51	1	.57	.48	.49	.57	.10	.21	.08	.03
55	2	.68	.42	.40	.07	.68	.05	.16	.04
58	2	.26	.37	.37	.14	.26	.08	.47	.05
65	3	.43	.41	.39	.07	.16	.43	.24	.09
70	1	.47	.53	.50	.47	.13	.18	.09	.13
72	3	.58	.55	.50	.10	.10	.58	.08	.15
73	3	.48	.51	.47	.10	.11	.48	.15	.16
75	2	.42	.46	.42	.12	.42	.09	.20	.17
79	3	.27	.24	.20	.18	.17	.27	.17	.20

Raw Scores

Mean = 9.02
S.D. = 3.54
KR #20 = 0.677

Corrected Scores

Mean = 5.74
S.D. = 4.45
KR #20 = 0.815

N = 5210

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

TABLE G-6

Item Data

Glaciers
Application

Item No.	Key	Diffi-culty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
24	2	.70	.33	.33	.03	.70	.16	.10	.01
28	4	.50	.33	.34	.18	.17	.14	.50	.01
30	4	.42	.29	.30	.07	.11	.40	.42	.01
31	3	.52	.40	.41	.11	.16	.52	.21	.01
32	3	.31	.30	.30	.22	.24	.31	.21	.02
39	4	.43	.50	.51	.38	.08	.10	.43	.01
41	2	.54	.24	.24	.13	.54	.18	.14	.01
43	1	.61	.56	.56	.61	.12	.12	.14	.01
44	2	.65	.57	.57	.10	.65	.09	.14	.02
45	2	.48	.46	.47	.22	.48	.11	.18	.02
46	2	.39	.39	.39	.16	.39	.20	.23	.02
52	2	.50	.41	.40	.21	.50	.13	.13	.04
53	2	.48	.41	.40	.14	.48	.13	.20	.05
54	1	.46	.47	.46	.46	.13	.17	.20	.05
56	1	.42	.47	.46	.42	.22	.15	.16	.05
57	1	.74	.47	.44	.74	.08	.06	.07	.05
62	1	.57	.55	.53	.57	.10	.12	.14	.07
76	3	.45	.37	.33	.06	.11	.45	.20	.18
77	1	37	.40	.37	.37	.11	.24	.10	.19
80	4	.27	.21	.18	.15	.17	.20	.27	.21

Raw Score		Corrected Scores
Mean = 9.80		Mean = 6.73
S.D. = 3.93		S.D. = 5.04
KR #20 = 0.731	N = 5210	KR #20 = 0.857

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

TABLE G-7

Item Data

Glaciers

Analysis

Item No.	Key	Difficulty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
23	3	.22	.23	.25	.49	.21	.22	.08	.01
26	1	.35	.29	.32	.35	.17	.19	.28	.01
33	4	.47	.25	.28	.11	.30	.11	.47	.01
34	2	.28	.10	.12	.17	.28	.44	.10	.01
35	3	.42	.31	.32	.05	.09	.42	.44	.01
36	1	.36	.28	.30	.36	.34	.16	.13	.01
38	4	.36	.30	.31	.20	.21	.22	.36	.01
48	3	.50	.36	.37	.08	.28	.50	.11	.02
59	3	.33	.20	.18	.09	.11	.33	.41	.06
60	1	.48	.47	.45	.48	.12	.21	.13	.06
61	4	.26	.21	.19	.30	.26	.11	.26	.07
63	4	.29	.31	.29	.16	.08	.40	.29	.08
64	3	.58	.53	.49	.06	.17	.58	.12	.08
66	4	.21	.18	.16	.12	.24	.34	.21	.09
67	1	.23	.27	.24	.23	.11	.12	.44	.10
68	2	.50	.52	.47	.19	.50	.10	.10	.11
69	4	.45	.46	.41	.06	.10	.27	.45	.13
71	2	.50	.46	.39	.16	.50	.10	.10	.14
74	4	.17	.28	.26	.10	.36	.20	.17	.17
78	2	.31	.37	.32	.20	.31	.19	.11	.20

Raw Scores		Corrected Scores	
Mean	= 7.25	Mean	= 3.45
S.D.	= 2.99	S.D.	= 3.64
KR #20	= 0.539	KR #20	= 0.706

N = 5210

*r_1 = Point biserial correlation, raw scores

r_2 = Point biserial correlation, corrected scores

APPENDIX C

Table of Contents

THE LISBON EARTHQUAKE

Directions

This is a test of your ability to apply skills and understandings which you have been learning since you entered school. You will be asked to read a passage called "The Lisbon Earthquake" and then you will be asked to answer questions about it. You may refer to the passage at any time during the test.

Your answer must be marked on the answer sheet which has been provided for you. Use only the special pencil that has been provided. Do not make any stray marks. If you make an error, erase it completely before marking your new answer.

Make no marks in the test booklet.

The following is a sample question to show you how your answers are to be marked. Study the sample and if you have any questions raise your hand.

SAMPLE | ON SEPARATE ANSWER SHEET

1. The article you will read is about the
   1. Lisbon Earthquake
   2. Northeast Floods
   3. Last days of World War II
   4. Spanish-American War

1. 1__ 2__ 3__ 4__ 5__

You should answer as many questions as you can. Do not spend a great amount of time on any one question. If you are in doubt about the answer to a question, then guess. Remember that you may refer to the reading passage at any time during the test.

R.P. Kropp and H. W. Stoker, Editors
Department of Educational Research and Testing
and Institute of Human Learning
Florida State University
Tallahassee

April, 1965

Experimental edition; to be used only for research purposes.

This experimental test was constructed as part of a research project supported by the U. S. Office of Education under Contract Number OE-4-10-019, Project Number, 2117.

THE LISBON EARTHQUAKE

Some catastrophes demand of man far more than relief and rehabilitation: they literally call for rethinking on a universal scale. This was so with the man-made disaster of Hiroshima. Similarly, the great earthquake at Lisbon on November 1, 1755, shook the minds of men.

While controversy surrounds most statistics dealing with the Lisbon earthquake, there is little doubt that it is one of the most severe recorded. Voltaire's classic description in the story, Candide, vividly paints the tragic scene after the earth started to tremble under the feet of the people of Lisbon: "The sea rose in foaming masses in the port and smashed the ships which rode at anchor. Whirlwinds of flame and ashes covered the streets and squares; the houses collapsed, the roofs were thrown upon the foundations, and the foundations were scattered; thirty-thousand inhabitants of every age and both sexes were crushed under the ruins."

In all, there were three shocks. The first, which lasted two minutes, shook the earth so slightly that an eyewitness recalled that he thought it had been caused by a passing vehicle. Two minutes later a second quake was felt, and this time its violence left no doubt as to what it was. During its ten minute visitation of terror, the dust from falling buildings was so great it obscured the sun. Next came another awful tremor, and the buildings which still remained standing now came tumbling down, bringing added dust, and plunging the city into total darkness. After twenty minutes of death-spelling noises, all became quiet. Then, to quote an eyewitness, "a very boisterous (stormy) wind" suddenly arose, fanning the flames of the candle-fed fires which had broken out all over the city.

Unfortunately, a combination of circumstances made the disaster greater than it might otherwise have been. For one thing, the quake occurred on All Saints' Day, which meant that candles had been burning since early morning in homes and churches. Then, to make matters worse, the earthquake struck at a bad time: shortly before ten in the morning--an hour when most of the people were at church. The violent movements of the earth caused the roofs of heavy stone to topple on the congregants, who, if they were not crushed to death, died in the flames.

The people experienced all the possible elements of horror. To falling stones and fires must be added the forty foot tidal wave which engulfed those who rushed to the quays after having escaped the earlier shocks. Furthermore, man, or at least a lower species, contributed looting and murder to the scene of despair. Valuable records, irreplaceable documents were lost, and, since there exists no inventory of Lisbon's art treasures of that time, we cannot even guess what the world has lost.

The older, medieval section of Europe's westernmost capital was completely destroyed. So, for that matter, were the towns within a distance of 20 leagues. "I write to you from the depths of the country," complained a survivor, "for there is not a habitable house left. Lisbon has vanished!" Built on a more substantial foundation of basalt, the newer section of Lisbon survived the earthquake.

The Lisbon earthquake, whose tremors were reportedly felt as far north as Norway and as far south as North Africa, made a profound impression on Europe. Great Britain was the first to offer help. Parliament voted the then tremendous sum of one hundred thousand pounds to aid the victims, in addition to gifts of food and clothing. Spain changed her tariff laws to favor Portugal's recovery. Also, large sums of money and provisions from all over Europe were generously offered by sympathetic nations and individuals.

Like today's moral and intellectual repercussions from man-made disastrous weapons, Lisbon's disaster registered severely on the mental seismographs of some of the outstanding thinkers of the eighteenth century. A noted historian of Portugal declares that to the little country on the Iberian Peninsula, the earthquake was "more than a cataclysm (disaster) of nature; it was a moral revolution."

So shattered was the moral and material structure of Lisbon society that it was seriously proposed that the government be transferred to Rio de Janeiro, the capital of its great colony! Fortunately, the crisis brought to complete power a ruthless, but exceedingly capable dictator, Pombal.

He was appointed Minister of Foreign Affairs and War by King Jose I in 1750 and quickly established himself as a dominant figure in Portuguese politics. The earthquake provided an opportunity for him to obtain complete power. On the day after the earthquake he told the Chief Justice to

appoint a special magistrate for each of the twelve wards of the city. These magistrates were given authority to carry out the government's emergency directives. Troops were rushed to Lisbon in order to maintain law and order and to assist in clearing up the ruins. Pombal's immediate concern was to prevent a plague; steps were taken to remove the bodies of men and animals from the ruins as quickly as possible, pools of stagnant waters were drained and contaminated food was destroyed. A most urgent matter was providing food and shelter for the survivors. Food centers were established and field kitchens were built. Prices of food and building materials were strictly controlled to prevent profiteering. Steps were taken to prevent looting. On November 4 immediate public execution after a summary trial was ordered for those caught looting the ruins.

Although many of Pombal's reforms were short lived, his great schemes and actual reforms shook Portuguese society loose from its medieval foundations. Starting with physical reconstruction while Lisbon was still smouldering, he built a new and more modern city. Temporary wooden structures were constructed outside the city to provide emergency housing and governmental offices. In early 1756 Pombal ordered unauthorized building in stone or brick stopped and the city was rebuilt according to a master plan. Taxation, civil law and public administration were reformed, new industries were set up, communications were improved, colonial relationships were re-evaluated, and education was revamped. Above all, by his ensuing power conflicts with the nobility and the clergy, Pombal helped Portugal advance on the road to a more modern society.

Meanwhile, elsewhere in Europe numerous accounts of the great earthquake were being published in virtually all languages. More than 20 reports, not including magazine articles, were published in 1755 in England alone! The great philosopher Immanuel Kant took time out from his studies to write a book on the theory of earthquakes. But the intellectual crisis in which Europe was embroiled for almost all the rest of the century took place mainly in France. Basically, the great quarrel of the age concerned the validity of the popular optimistic philosophy (hopeful outlook) of Leibniz, who believed that "What is, is Right," and that this is the "best of all possible worlds."

Leibniz stated that man could have no free will in a perfect world and that "Our world is suited to our desires and appetites." He believed that the world was built on a plan which harmonizes with the moral government

of its inhabitants and theorized that the past, present, and future have already been set with as much order and harmony as possible. Leibniz surmised that "the world must be destroyed and repaired by natural means, at such times as the government of spirits may demand it for the punishment of some and the reward of others." He felt that evil tends to evoke a greater good in the long run and maintained, "It is impossible to make the world better than it is, not only as a whole and in general, but also for ourselves in particular."

Voltaire, in his long poem, "The Lisbon Earthquake," vigorously attacked the Leibniz philosophy. He regarded it as unprogressive in that "physical evil deserved man's attention." It was also a cruel dogma, he believed, in that it implied that "your particular misfortune is nothing; it contributes to the universal good." Voltaire expressed faith in progress which, he said, depended upon the good sense of mankind.

Leibniz, however, held that we should be content with the order of the past because it is in conformity with the absolute will of God. Although Leibniz suggested that we should make the future in conformity with the presumed will of God, he cautioned against becoming upset if we were unsuccessful.

Rousseau, in an impassioned refutation (answer), maintained an "all is good" theme. Man must be patient and recognize evil as the consequence of his own nature. Furthermore Rousseau claimed that civilization had corrupted man. Although Rousseau looked to the past and said progress was an illusion, he was later to expound, in his Social Contract, a theory of rule by the consent of the governed and actually advocated revolt by the people if they were unfairly ruled.

In Candide Voltaire, as we know, returned to the fray with slashing attacks on Rousseau and Leibniz for their views concerning human progress. Practically all the philosophers of the eighteenth century took sides in what has been called the "theology of earthquakes." Such was the exchange of arguments, in fact, that the wordy Dr. Johnson complained that he was weary of hearing about the subject.

While no such clear-cut philosophical discussion fills our twentieth century air, we scarcely need be reminded that, once again, recent catastrophes have sent man to meditate on life's eternal questions. Obviously, man is worried about possible misuse of fission and fusion.

In addition, Nature, with her unlady-like hurricanes of recent years, and the devastating floods of the past summer, has intruded into what had begun to seem to many like a man-manipulated world. While we are, today, better equipped for relief and rehabilitation than the Portuguese were two hundred years ago, it is well to remember that as in the case of the Lisbon disaster, the Northeast floods were not even predicted, much less staved off.

Nature's calamities and their aftermath of re-evaluation are still very much with us.

Adapted from

Albert Alexander, "The Lisbon Earthquake," Social Education, Vol. XX, No. 1, January, 1956, pp. 27-28.

and

An adaptation of pages 74-85 from THE LISBON EARTHQUAKE by T.D. Kendrick. Copyright (C) 1957 by T.D. Kendrick. Published by J.B. Lippincott Company.

1. The disasters at Hiroshima and Lisbon were alike in that both
   (K) 1. happened in the spring.
   2. were man-made.
   3. called for universal re-thinking.
   4. were earthquakes.

2. Who was the author of <u>Candide</u>?
   (K) 1. Kant
   2. Leibniz
   3. Rousseau
   4. Voltaire

3. What was Voltaire's profession?
   (K) 1. orator
   2. writer
   3. painter
   4. physician

4. How many earth shocks were felt in the Lisbon earthquake?
   (K) 1. 2
   2. 3
   3. 4
   4. 5

5. How long was the first earthquake shock at Lisbon?
   (K) 1. two minutes
   2. ten minutes
   3. twenty minutes
   4. thirty-two minutes

6. On which of the following days did the Lisbon earthquake occur?
   (K) 1. All Saints' Day
   2. Christmas
   3. Easter
   4. Good Friday

7. Which of the following <u>was not</u> true of the Lisbon earthquake?
   (K) 1. It happened during evening services.
   2. A tidal wave also struck.
   3. The sun was obscured part of the time.
   4. It occurred about 10 o'clock in the morning.

8. The newest section of Lisbon survived the earthquake because
   (K) 1. it was the most westerly part of the city.
   2. there was only one tremor in that section.
   3. it was built on a basalt foundation.
   4. there was no fire in that section.

9. The first country to offer aid to Lisbon was
   (K) 1. France.
   2. Great Britain.
   3. Norway.
   4. Spain.

10. Parliament is a representative assembly in
   (K) 1. Great Britain.
   2. Norway.
   3. Portugal.
   4. Spain.

11. Which of the following <u>was not</u> offered by Great Britain to help Portugal?
   (K) 1. clothing
   2. food
   3. lowering of trade tariffs
   4. money

12. King Jose I appointed Pombal to the position of
   (K) 1. Minister of Foreign Affairs and War.
   2. Chief Justice.
   3. Dictator.
   4. Chief Magistrate.

13. Before the earthquake, Pombal was
   (K) 1. dictator of Portugal.
   2. a dominant figure in Portuguese politics.
   3. influential among the common people.
   4. a noted philosopher.

14. Immediately following the earthquake Pombal's major concern was to
   (K) 1. rebuild the city.
   2. prevent looting.
   3. provide shelter.
   4. prevent a plague.

15. Prices of food and building material were strictly controlled after the earthquake to prevent
(K) 1. inflation.
2. profiteering.
3. shortages.
4. waste.

16. Persons caught looting the ruins were punished of execution
(K) 1. without a fair trial.
2. by the soldiers who caught them.
3. after a summary trial.
4. after a long delay.

17. Lisbon was rebuilt
(K) 1. in a haphazard manner.
2. outside the old city limits.
3. without using brick or stone.
4. according to a master plan.

18. Voltaire's poem, "The Lisbon Earthquake," was an attack on
(K) 1. Pombal's emergency directives.
2. Leibniz's philosophy.
3. Kant's philosophy.
4. King Jose I's lack of action.

19. Rousseau considered that the misfortunes resulting from the earthquake were nothing compared with the good which ultimately resulted from it. Which one of the following agreed with him?
(K) 1. Johnson
2. Kant
3. Leibniz
4. Voltaire

20. Order the following statements according to the time of occurrence.
(K) A. Countries offered goods or concessions to aid Portugal's recovery.
B. A severe earthquake destroyed much of the capital of Portugal in 1755.
C. A great deal of moral and intellectual discussion took place concerning the earthquake.
D. The rebuilding of Portugal, especially of the city of Lisbon, went hand in hand with a reform movement.

The correct order is
1. B, A, D, C
2. B, D, A, C
3. B, C, D, A
4. B, D, C, A

21. Dr. Johnson said that he was "weary of hearing about the subject." He was dissatisfied with the
(C) 1. lack of interest in the Lisbon problem.
2. overly optimistic viewpoint of others.
3. lack of interest in philosophical arguments.
4. amount of material written about the subject.

22. Of the statements below, the one which means the <u>opposite</u> of "This is the best of all possible worlds" is
(C) 1. "progress is our most important goal."
2. "let sleeping dogs lie."
3. "better safe than sorry."
4. "moderation in all things is to be desired."

23. Why was the Lisbon earthquake the basis of much philosophical discussion?
(An) 1. A capital of a country was destroyed.
2. Many of Portugal's treasures were lost.
3. Pombal's form of government was a dictatorship.
4. 30,000 or more inhabitants were killed in the disaster.

24. The response of the English Parliament in 1775 served the function of what agency today?
(Ap) 1. United Nations
2. International Workers of the World
3. Red Cross
4. Overseas Refugee Association

25. The article states that the tremors of the Lisbon earthquake were felt over a wide area. In which of the following countries were these tremors most violent?
(C) 1. England
2. Germany
3. Italy
4. Spain

26. If the earthquake had not occurred, which of the following would have been most likely?
(An) 1. King Jose I would have become an absolute monarch.
2. Pombal would have lost his power and influence.
3. Pombal would have had a less profound influence on Portugal's future.
4. The Chief Justice would have become the most powerful person in Portugal.

27. Which of the following words might best describe basalt?
(C) 1. brittle
2. sandy
3. solid
4. spongy

28. What is the relationship between the following statements?
(Ap) A. Pombal's conflict with the nobility and the clergy.
B. Modernization of Portugal.
1. A caused by B
2. B caused by A
3. A and B are related, but one did not cause the other.
4. A and B are unrelated.

29. "Man or at least a lower species contributed looting and murder to the scene of despair." In this sentence "lower species" most nearly means
(C) 1. non-noblemen and working men.
2. looters and robbers.
3. animals of high order.
4. morally inferior men.

30. What is the relationship between the following statements?
(Ap) A. Charging of unreasonable prices for rent.
B. Rent control following the earthquake.
1. A caused B.
2. B caused A.
3. A and B are related, but one did not cause the other.
4. A and B are unrelated.

31. What is the relationship between the following statements?
(Ap) A. Many ships were destroyed during the earthquake.
B. Tariff regulations were changed following the earthquake.
1. A caused B.
2. B caused A.
3. A and B are related, but one did not cause the other.
4. A and B are unrelated.

32. Which of the following philosophers would have most likely supported Pombal's policies?
(Ap) 1. Kant
2. Leibniz
3. Rousseau
4. Voltaire

* * * * * * * * * *

<u>NOTE</u>: The following information <u>must</u> be used in answering questions 33-35:

Voltaire and Leibniz might have argued about whether or not the following conclusion was true: "<u>Great social, political, and economic reforms are needed and men are justified in putting forth their greatest effort in this direction.</u>" Decide how Voltaire and Leibniz would have used each of the following statements, if at all, in arguing about this conclusion. Mark your answer according to the following key:

<u>KEY</u>:
1. Voltaire would have used the statement to argue that the conclusion was true.
2. Leibniz would have used the statement to argue that the conclusion was true.
3. The statement has no bearing on the argument.
4. The statement relates to the argument but neither would have used it.

33. Evil will bring about a greater good.
(An)

34. All events have been predetermined since the beginning of time.
(An)

35. Miserable conditions should not be tolerated because man has enough common sense to overcome them.
(An)

* * * * * * * * * *

36. If <u>man</u> were to misuse nuclear fission and fusion which person's philosophy would have the least relevance?
(An)
1. Leibniz
2. Pombal
3. Rousseau
4. Voltaire

37. Which one of the following conditions would have resulted in the reduction of damage to the older section of Lisbon?
    (C) 1. The existence of a basalt foundation under all of Lisbon.
    2. The occurrence of the earthquake on a non-religious holiday.
    3. The use of wooden roofs for buildings.
    4. The location of the center of the earthquake ten leagues away.

38. Which one of the following would most likely not have been a "good Samaritan"?
    (An) 1. Kant
    2. Leibniz
    3. Pombal
    4. Voltaire

39. What is the relationship between the following statements?
    (Ap) A. Voltaire's poem "The Lisbon Earthquake."
    B. Leibniz's philosophy.
    1. A caused B.
    2. B caused A.
    3. A and B are related, but one did not cause the other.
    4. A and B are unrelated.

40. The number of persons killed in the earthquake was undoubtedly increased because stone was used for
    (C) 1. sea walls.
    2. streets and sidewalks.
    3. sidewalls of buildings.
    4. roofs of buildings.

41. What is the relationship between the following statements?
    (Ap) A. Scarcity of food.
    B. Scarcity of building materials.
    1. A caused B.
    2. B caused A.
    3. A and B are related, but one did not cause the other.
    4. A and B are unrelated.

42. A viewpoint which cannot be found in the reading passage is that of
(C)
1. an observer.
2. a scientist.
3. a philosopher.
4. an historian.

43. In what way were the results of the destruction of Hiroshima similar to those of the Lisbon earthquake?
(Ap)
1. What Pombal did for Lisbon, the U.S. military government did for Hiroshima.
2. The disaster led to philosophical arguments around the world.
3. Hirohito strengthened his position in Japanese government and used dictatorial powers to help his people.
4. The large nations of the world came to the aid of the citizens of Hiroshima.

44. If you lived in a country where most of the citizens were poor and lived in slums and the rulers were rich and lived in palaces, what would you do if you believed in the later teachings of Rousseau?
(Ap)
1. Urge the people to revolt against the rulers.
2. Urge your fellow citizens to let well enough alone.
3. Urge the government to build schools for the poor.
4. Remind your fellow citizens that progress is bound to occur.

45. Which of the following statements best represents Pombal's philosophy of life?
(Ap)
1. What is to be will be.
2. Might makes right.
3. God punishes guilty and innocent alike.
4. Bury the dead and feed the living.

46. If an earthquake of the same magnitude as the Lisbon earthquake were to occur today in a large U.S. city,
(Ap)
1. aid would be needed from foreign countries.
2. the city would be self-supporting, and outside aid would be unnecessary.
3. a philosophical argument similar to Lisbon's would begin.
4. it would make headlines for a few weeks, then it would be forgotten.

47. "What is, is right," is most nearly equivalent to
(C)
1. "what is to be, will be."
2. "the end justified the means."
3. "might makes right."
4. "the sky is the limit."

48. What characteristic of a medieval society discouraged modernization of Portugal?
(An)
1. Political and economic power of the nobles and clergy.
2. The existence of an old section in the towns.
3. Great emphasis on religion.
4. Lack of trade with other countries.

49. At the time of the earthquake, Lisbon society could best be described as
(C)
1. enlightened.
2. medieval.
3. progressive.
4. modern.

50. Why were most of the discussions of the Lisbon earthquake philosophical?
(C)
1. Few written accounts were available.
2. There were few survivors.
3. There was no accurate means of describing the disaster.
4. Philosophers were the spokesmen of the time.

51. The statement, "Lisbon is the capital of Portugal," is
(C)
1. a guess.
2. not true.
3. supported by the reading passage.
4. impossible to make from the reading selection alone.

52. Which of the following statements is best supported by the philosophy of Voltaire?
(Ap)
1. Social welfare programs should be curtailed.
2. The American foreign aid program should be eliminated.
3. People with children in school should pay a tax to support education.
4. Big cities should start slum clearance projects.

53. Which of the following would most likely have done as Nero did--"Fiddle while Rome burned"?
(Ap) 1. Kant
2. Pombal
3. Rousseau
4. Voltaire

54. What action, if any, would Pombal probably have taken if most of the skilled craftsmen like carpenters and bricklayers had left Lisbon immediately following the earthquake?
(Ap) 1. None, because there would have been fewer people to feed and shelter.
2. None, because other areas needed these people more than Lisbon did.
3. Action to return them so they could help rebuild the city.
4. Action to return them so they could be punished for leaving.

55. A "mental seismograph" is a
(C) 1. scientific device for detecting ideas.
2. figure of speech for the mind.
3. mental record.
4. mechanical device for recording earthquakes.

56. What is the relationship between the following statements?
(Ap) A. There were many fires in Lisbon after the earthquake.
B. Most of the inhabitants of Lisbon observed religious holidays.
1. A caused B.
2. B caused A.
3. A and B are related, but one did not cause the other.
4. A and B are unrelated.

57. What is the relationship between the following statements?
(Ap) A. Scarcity of food.
B. Control of prices of food.
1. A caused B.
2. B caused A.
3. A and B are related, but one did not cause the other.
4. A and B are unrelated.

58. How could you best describe the statement that "Lisbon has vanished? It is
(C) 1. absurd
2. accurate
3. exaggerated
4. unsubstantiated

* * * * * * * *

NOTE: The following information must be used in answering questions 59-61:

Voltaire and Leibniz might have argued about whether or not the following conclusion was true: Great social, political, and economic reforms are needed and men are justified in putting forth their greatest effort in this direction.". Decide how Voltaire and Leibniz would have used each of the following statements, if at all, in arguing about this conclusion. Mark your answer according to the following key:

KEY: 1. Voltaire would have used the statement to argue that the conclusion was true.
2. Leibniz would have used the statement to argue that the conclusion was true.
3. The statement has no bearing on the argument.
4. The statement relates to the argument, but neither would have used it.

59. History shows that civilizations improve over time.
(An)

60. Man has a responsibility to society.
(An)

61. Natural disasters cause great suffering.
(An)

* * * * * * * *

62. What is the relationship between the following statements?
(Ap) A. An earthquake
B. The rise of a dictator
1. A caused B
2. B caused A
3. A and B are related, but one did not cause the other.
4. A and B are unrelated.

63. Which one of the following statements could be attributed to Voltaire?
(An) 1. Research is more important than application.
2. Benefit to man is the primary goal of science.
3. Science serves the purpose of discovering the natural harmony and order of the world.
4. The study of science is not a proper activity for man.

64. Which fact about Pombal is most consistent with the belief that he accepted the philosophy of Voltaire?
(An) 1. He did not move the capital to Rio de Janeiro.
2. He was ruthless.
3. He broke with tradition.
4. He used Voltaire's poem, "The Lisbon Earthquake," as his guide to reconstruct Lisbon.

65. Pombal was able to assume complete power following the earthquake because he was
(C) 1. a member of the nobility.
2. the spokesman of the people.
3. established in the power structure of Portugal.
4. well liked by the king and his court.

66. Which of the following best describes the reason aid was given to Lisbon?
(An) 1. A feeling of sorrow and charity.
2. A desire for extra rights in rich colonies.
3. A feeling of loyalty and friendship for a neighboring country.
4. a desire to establish good relations for possible future benefit.

67. Which of the following is both a sequential relationship and a cause-and-effect relationship?
(An) 1. All Saints' Day celebration--Lisbon earthquake.
2. Consideration of moving to Rio de Janeiro--Pombal's rise to power.
3. "Theology of Earthquakes"--extensive news coverage of the Lisbon earthquake.
4. The Lisbon earthquake--European aid to Portugal.

68. Assume that flood waters have ruined the agricultural area of a foreign country and millions will starve unless aid is received. Were he living now, which one of the following individuals would most likely advocate that U.S. surplus crops be made available to the distressed country?
(An) 1. Kant
2. Leibniz
3. Rousseau
4. Voltaire

69. Many of Pombal's measures were short lived because
(An) 1. they were too modern.
2. they were emergency reforms.
3. they were not endorsed by the clergy.
4. of foreign intervention in Portugal's affairs.

70. In which country is the westernmost European capital?
(C) 1. France
2. Norway
3. Portugal
4. Spain

71. Which one of the following contributed most toward the continuing discussion of the "Theology of Earthquakes"?
(An) 1. The issue was very controversial and many philosophers were interested in it.
2. The progress of science was at stake for the rest of the century.
3. Acceptance or rejection of Leibniz's philosophy would govern man's attitude toward his world.
4. The common man was interested in having a better life.

72. The greatest damage to articles such as books, tapestries, and paintings was probably caused by
(C) 1. fire.
2. tremors.
3. water.
4. wind.

73. From which point of view was the passage written?
(C) 1. historical
2. military
3. political
4. scientific

74. Which one of the following would have been most likely to break with tradition?
(An)
1. Kant
2. Leibniz
3. Rousseau
4. Voltaire

75. The article states that there were four primary causes of death in the Lisbon earthquake. Which one probably took the fewest lives?
(C)
1. falling objects
2. fire
3. murder
4. tidal wave

76. What is the relationship between the following statements?
(Ap)
A. A philosophical controversy which lasted the better part of a century.
B. The Lisbon earthquake.
   1. A caused B
   2. B caused A
   3. A and B are related, but one did not cause the other.
   4. A and B are unrelated.

77. What is the relationship between the following statements?
(Ap)
A. Spain changed its tariff laws.
B. Pombal rose to power as a dictator.
   1. A caused B
   2. B caused A
   3. A and B are related, but one did not cause the other.
   4. A and B are unrelated.

78. Which one of the following conclusions is <u>accurately</u> based upon some part of the reading passage?
(An)
1. The earthquake was on Sunday.
2. The Lisbon police were ineffective in maintaining law and order.
3. The English were the only generous people in the 18th century Europe.
4. The damage would have been just as bad on any other popular religious holiday.

79. If the Lisbon earthquake had happened after Thomas Edison invented the light bulb, which of the following would have been most likely?
(C)
1. The tidal wave could have been averted.
2. There would have been fewer fires.
3. There would have been a minimum amount of dust.
4. The "boisterous," stormy wind would not have arisen.

80. What is the relationship between the following statements?
(Ap)
A. Voltaire, Rousseau, and other prominent philosophers were Frenchmen.
B. The center of controversy regarding the "Theology of Earthquakes" was in France.
1. A caused B
2. B caused A
3. A and B are related, but one did not cause the other.
4. A and B are unrelated.

THE LISBON EARTHQUAKE
(Booklet Number 2)

Directions

Your answers to each of the items 81-95 should be written directly on the test booklet in the space provided below each item. Should you need additional space for an item write "continued" then write the remainder of your answers on the back of the last page of the booklet.

You received with this test booklet another booklet which contains the article about the Lisbon earthquake. You may refer to it any time during the examination. You might wish to scan the article after you read these directions to refresh your memory.

Your answers to these items will depend on your understanding and thoughts about the article. Each of you might write an entirely different answer to an item because there are many possible answers to each of them. So please write whatever you believe to be appropriate.

Answers to these items do not appear as such in the article. Therefore, you will certainly answer incorrectly if you copy material directly from the article.

81. (Sy) Suppose that you had held an influential position in Lisbon and that you were opposed to moving the capital to Rio de Janeiro. Other influential people supported the proposed move. Briefly outline the reasons you would have used in an attempt to convince them that the capital should remain in Lisbon.

82. (Sy) Suppose a political cartoonist wished to make the philosophy of Leibniz appear foolish. Draw or describe a cartoon which would accomplish the above purpose.

83. (Sy) The article implies that the nobility opposed Pombal's reforms. Suppose you were a member of the Portuguese nobility. In a few sentences describe why you would have opposed Pombal.

84. (Sy) Events are described differently by different people according to the views which they hold. Describe how a Lisbon priest might have described the earthquake and its immediate aftermath.

85. (Sy) The Lisbon Earthquake article is approximately three pages in length. Attempt to summarize the entire article in five sentences.

86-95. Suppose that you have been asked to evaluate the author's qualifications to write an article on the Lisbon earthquake. Listed below are several statements which could be true about the author. Assume each statement to be true. Do not consider any relationships which might exist between the statements.

"qualified"--if the statement leads you to believe the author is qualified to write about the Lisbon earthquake.

"not qualified"--if you believe the statement leads you to believe the author is not qualified to write about the Lisbon earthquake.

"no effect"--if you believe the statement has no bearing on the author's qualifications, or lack of them, to write about the Lisbon earthquake.

In the space provided below each statement write a brief statement giving the reason(s) why you marked the answer you did.

86. He wrote a book condemning Leibniz.
(Ev) qualified ____ not qualified ____ no effect ____

87. He is a professor of world history at a large university.
(Ev) qualified ____ not qualified ____ no effect ____

88. He is a retired junior high school librarian.
(Ev) qualified ____ not qualified ____ no effect ____

89. He has never been outside the United States.
(Ev) qualified ____ not qualified ____ no effect ____

90. He lived in Portugal until he was 25 years old.
(Ev) qualified ____ not qualified ____ no effect ____

91. He wrote this article while teaching social studies
(Ev) in a small town.
qualified ____ not qualified ____ no effect ____

92. He is a well known philosopher and has written many
(Ev) books and articles.
qualified ____ not qualified ____ no effect ____

93. He is a well known author of historical fiction.
(Ev)
qualified ____ not qualified ____ no effect ____

94. Pombal and he were good friends.
(Ev) qualified ____ not qualified ____ no effect ____

95. He has been blind since birth.
(Ev) qualified ____ not qualified ____ no effect ____

TABLE L-1

LISBON EARTHQUAKE

Percentiles*

Raw Score	Knowledge 9	Knowledge 10	Knowledge 11	Knowledge 12	Comprehension 9	Comprehension 10	Comprehension 11	Comprehension 12
20	93	86	82	80		99	99	99
19	77	66	61	55	99	97	96	95
18	64	47	42	34	98	94	91	88
17	48	33	27	20	96	90	85	79
16	35	23	18	11	92	83	77	69
15	27	17	12	08	87	75	66	60
14	20	12	09	05	80	66	56	48
13	15	09	06	03	71	57	47	39
12	11	06	04		60	47	39	29
11	08	04	03	02	49	38	30	21
10	06	03	02		40	30	21	14
9	04				30	22	16	09
8	03	02		01	22	15	11	06
7			01		16	10	08	03
6	01				10	06	05	02
5		01			06	04	03	01
4					03	02		
3					01	01	01	
2								
1								
0								

*Percentiles computed on the per cent of students scoring below a given raw score. Percentiles for this form were based on the following numbers of students:

Grade 9 - 1194
Grade 10 - 1451
Grade 11 - 1298
Grade 12 - 1192

TABLE L-2

LISBON EARTHQUAKE

Percentiles

Raw Score	Application 9	Application 10	Application 11	Application 12	Analysis 9	Analysis 10	Analysis 11	Analysis 12
20								
19								
18		99	98	98				
17	99	97	96	94			99	99
16	98	94	92	88	99	98	97	96
15	96	88	86	81	98	96	94	91
14	92	81	77	72	97	92	90	84
13	88	74	68	59	94	87	83	76
12	81	65	57	47	92	81	76	68
11	73	55	47	38	88	75	68	60
10	62	46	37	28	84	68	60	52
9	52	36	27	20	74	60	51	42
8	40	28	20	14	62	51	41	32
7	29	19	14	09	51	39	30	22
6	18	12	10	06	38	26	21	15
5	11	07	05	03	24	16	14	08
4	04	04	03		13	09	06	04
3	02	02	02	01	06	04	04	02
2	01	01			02	02	01	
1			01			01		
0								

TABLE L-3

LISBON EARTHQUAKE

Percentiles

Raw Score	Synthesis 9	Synthesis 10	Synthesis 11	Synthesis 12	Evaluation 9	Evaluation 10	Evaluation 11	Evaluation 12
20								
19								
18								
17								
16								
15								
14								
13								
12								
11			99					
10		99	98	99				99
9		98	97	98		99	99	98
8		97	95	96		98	97	97
7	99	94	92	92	99	96	95	94
6	97	89	87	85	96	93	91	90
5	95	82	80	71	94	90	86	84
4	87	71	69	56	88	84	77	75
3	75	56	52	40	81	76	67	64
2	58	40	34	26	69	66	54	51
1	36	25	20	17	53	52	41	36
0	17	14	11	08	34	36	27	23

PRECEDING PAGE BLANK-NOT FILMED

TABLE L-4

Item Data

Lisbon Earthquake
Knowledge

Item No.	Key	Difficulty	r_1*	r_2 *	Proportion Choosing Each Response 1	2	3	4	Omit
1	3	.78	.50	.50	.02	.09	.78	.11	.00
2	4	.93	.47	.47	.02	.03	.02	.93	.00
3	2	.78	.40	.40	.04	.78	.15	.03	.00
4	2	.84	.33	.33	.04	.84	.11	.01	.00
5	1	.94	.35	.34	.94	.04	.01	.00	.00
6	1	.99	.36	.35	.99	.00	.01	.01	.00
7	1	.90	.42	.41	.90	.03	.03	.03	.00
8	3	.91	.47	.46	.03	.03	.91	.02	.00
9	2	.90	.44	.43	.05	.90	.02	.03	.01
10	1	.91	.46	.46	.91	.02	.05	.03	.00
11	3	.87	.49	.48	.03	.01	.87	.09	.01
12	1	.86	.52	.51	.86	.03	.04	.06	.00
13	2	.50	.32	.32	.35	.50	.09	.05	.01
14	4	.83	.55	.55	.06	.06	.05	.83	.00
15	2	.90	.50	.50	.04	.90	.04	.02	.00
16	3	.89	.46	.46	.03	.07	.89	.01	.00
17	4	.86	.49	.47	.02	.07	.04	.86	.00
18	2	.89	.56	.54	.03	.89	.04	.02	.00
19	3	.68	.44	.43	.07	.12	.68	.13	.00
20	1	.81	.47	.46	.81	.07	.08	.03	.01

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

Raw Scores		Corrected Scores	
Mean =	16.96	Mean	15.95
S.D. =	2.91	S.D.	3.85
KR #20 =	0.758	KR #20 =	0.885

N = 5188

TABLE L-5

Item Data

Lisbon Earthquake
Comprehension

Item No.	Key	Difficulty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
21	4	.42	.43	.44	.06	.30	.22	.42	.01
22	1	.35	.39	.40	.35	.24	.12	.28	.01
25	4	.67	.42	.42	.15	.07	.10	.67	.01
27	3	.77	.36	.37	.07	.06	.77	.09	.00
29	4	.45	.42	.43	.08	.40	.07	.45	.01
37	1	.72	.38	.38	.72	.12	.07	.08	.01
40	4	.75	.37	.38	.03	.03	.18	.75	.01
42	2	.53	.43	.44	.18	.53	.06	.22	.01
47	1	.63	.27	.27	.63	.22	.09	.05	.01
49	2	.74	.47	.48	.07	.74	.15	.03	.01
50	4	.54	.45	.46	.14	.13	.17	.54	.02
51	3	.61	.36	.36	.05	.11	.61	.21	.02
55	2	.31	.32	.32	.10	.31	.35	.22	.03
58	3	.51	.38	.38	.09	.22	.51	.15	.03
65	3	.59	.45	.44	.08	.14	.59	.15	.05
70	3	.48	.46	.44	.15	.17	.48	.12	.07
72	1	.77	.52	.47	.77	.05	.08	.03	.08
73	1	.59	.45	.43	.59	.03	.18	.11	.08
75	3	.79	.55	.50	.04	.04	.79	.04	.09
79	2	.75	.51	.47	.05	.75	.07	.03	.11

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

Raw Score

Mean = 11.96
S.D. = 3.87
KR #20 = 0.743

Corrected Score

Mean = 9.48
S.D. = 4.94
KR #20 = 0.862

N = 5188

TABLE L-6

Item Data

Lisbon Earthquake
Application

Item No.	Key	Difficulty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
24	3	.47	.38	.39	.29	.10	.47	.13	.01
28	2	.21	.16	.16	.30	.21	.28	.21	.01
30	1	.39	.36	.36	.39	.17	.26	.18	.01
31	4	.44	.26	.27	.24	.05	.27	.44	.00
32	4	.41	.42	.42	.15	.24	.19	.41	.01
39	2	.47	.44	.44	.14	.47	.24	.15	.01
41	3	.68	.31	.30	.04	.05	.68	.23	.01
43	2	.56	.44	.44	.14	.56	.10	.19	.01
44	1	.50	.35	.35	.50	.19	.11	.19	.01
45	4	.55	.42	.42	.17	.19	.07	.55	.01
46	4	.40	.34	.34	.17	.13	.29	.40	.02
52	4	.46	.44	.44	.25	.12	.14	.46	.03
53	3	.52	.40	.40	.20	.12	.52	.14	.02
54	3	.75	.45	.44	.08	.05	.75	.10	.02
56	2	.51	.44	.44	.05	.51	.16	.25	.03
57	1	.74	.46	.44	.74	.09	.11	.04	.03
62	1	.66	.46	.44	.66	.06	.10	.14	.04
76	2	.68	.50	.47	.06	.68	.09	.08	.09
77	4	.40	.28	.26	.05	.09	.36	.40	.10
80	1	.33	.37	.36	.33	.10	.30	.16	.11

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

Raw Scores		Corrected Scores	
Mean	= 10.12	Mean	= 7.00
S.D.	= 3.66	S.D.	= 4.72
KR #20	= 0.689	KR #20	= 0.834

N = 5188

TABLE L-7

Item Data

Lisbon Earthquake
Analysis

Item No.	Key	Difficulty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
23	4	.44	.29	.30	.18	.23	.14	.44	.01
26	3	.69	.43	.43	.13	.12	.69	.06	.01
33	2	.42	.25	.25	.23	.42	.21	.13	.62
34	2	.41	.27	.27	.18	.41	.23	.17	.02
35	1	.48	.50	.50	.48	.17	.15	.18	.02
36	1	.26	.24	.24	.26	.21	.25	.27	.01
38	2	.41	.43	.43	.17	.41	.24	.16	.01
48	1	.58	.43	.43	.58	.17	.15	.09	.02
59	4	.12	.16	.16	.36	.33	.15	.12	.04
60	1	.46	.50	.49	.46	.24	.15	.11	.05
61	3	.32	.33	.31	.20	.20	.32	.24	.05
63	2	.48	.47	.46	.11	.48	.29	.08	.05
64	3	.38	.47	.46	.22	.14	.38	.22	.05
66	1	.37	.39	.38	.37	.05	.33	.20	.05
67	4	.52	.50	.48	.21	.10	.12	.52	.06
68	4	.46	.57	.55	.12	.18	.18	.46	.06
69	2	.49	.41	.38	.17	.49	.16	.11	.07
71	3	.21	.06	.04	.46	.13	.21	.12	.08
74	4	.38	.54	.52	.14	.19	.20	.38	.09
78	4	.37	.25	.23	.26	.18	.08	.37	.11

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

Raw Scores		Corrected Scores	
Mean	= 8.22	Mean	= 4.56
S.D.	= 3.60	S.D.	= 4.58
KR #20	= 0.684	KR #20	= 0.824

N = 5188

APPENDIX D

Table of Contents

STAGES OF ECONOMIC GROWTH

Directions

This is a test of your ability to apply skills and understandings which you have been learning since you entered school. You will be asked to read a passage called "Stages of Economic Growth" and then you will be asked to answer questions about it. You may refer to the passage any time during the test.

Your answer must be marked on the answer sheet which has been provided for you. Use only the special pencil that has been provided. Do not make any stray marks. If you make an error, erase it completely before marking your new answer.

Make no marks in the test booklet.

The following is a sample question to show you how your answers are to be marked. Study the sample and if you have any questions raise your hand.

SAMPLE | ON SEPARATE ANSWER SHEET

1. The article you will read is about
   1. the stages of economic growth
   2. the Boxer Rebellion
   3. The Seven Years War
   4. The Spanish-American War

1. 1___ 2___ 3___ 4___ 5___

You should answer as many questions as you can. Do not spend a great amount of time on any one question. If you are in doubt about the answer to a question, then guess. Remember that you may refer to the reading passage at any time during the test.

R.P. Kropp and H. W. Stoker, Editors
Department of Educational Research and Testing
and Institute of Human Learning
Florida State University
Tallahassee

April, 1965.

Experimental edition; to be used for research purposes.

EG-2

This experimental test was constructed as a part of a research project supported by the U.S. Office of Education under Contract Number OE-4-10-019, Project Number 2117.

STAGES OF ECONOMIC GROWTH

In several magazine articles and in one book, W.W. Rostow stated that it is possible to identify each nation, with respect to its economic development, as lying in one of five categories:

(1) The traditional society
(2) The precondition for take-off
(3) The take-off
(4) The drive to technological maturity
(5) The age of high mass-consumption

The basic principle of his theory is that, at any given time in an economy, the rapid rate of growth in a relatively few leading industries contributes toward maintaining the over-all strength of that economy. Rostow considers economic change to be a result of political and social as well as economic forces. Pursuing this thought further, Rostow quotes Keynes' dictum: "If human nature felt no temptation to take a chance, no satisfaction (profit apart) in constructing a factory, a railway, a mine or a farm, there might not be much investment merely as a result of cold calculation."

The Traditional Society

The main economic fact about the first stage, traditional society, is the existence of a ceiling on the level of attainable production per head. This ceiling stems from the fact that the potentialities which flow from modern science and technology either are not available or are not applied in a regular fashion. Traditional societies undergo constant change in production due to harvests, plagues, discoveries of new crops and so on. Varying degrees of manufacture develop and agricultural activity rises with improvements like irrigation, but production is still limited by the inaccessibility of modern science and the lack of a systematic understanding of the physical environment capable of making invention a regular flow.

The traditional society is basically agricultural with food production typically absorbing 75% or more of the working force. From this situation follows a social structure which thwarts a man's attempts toward improving his lot in life. Wealth and power are concentrated in the hands of those who control the land, with the real political

power tending to lie in the regions rather than in the central government. Clan and family ties play a significant role.

The Pre-Condition for Take-off (Transitional Period)

The second state, the precondition for take-off, is also referred to as the transitional period. Usually, this period begins as a result of aggression by more advanced societies. Essentially, the difference between the traditional society and a more modern society is related to the rate of investment. The traditional society's rate of investment is low (under 5% national income) in comparison to its rate of population increase.

To get the rate of investment up, three sectors--agriculture, export, and social overhead--of the economy are particularly important.

Agriculture--An increased food supply is required to meet the likely rise in population and the growing urban population. Agriculture must help meet the foreign exchange bill for capital development. This can be done directly by selling surplus abroad or indirectly be reducing food imports. Rising farm income must furnish taxes to finance governmental functions and farm surplus income must be controlled by men who will invest in trade and industry and who will reinvest their profits as productivity rises.

Exports--Exports can provide a quick source of money for investment in industry. It takes time for industry to gather strength and there are big bills to pay; therefore, a good part of the investment money must come from rapid increases in production and exportation. Quick-yielding changes in productivity can most readily be applied to the extraction and processing of natural resources.

Social overhead capital--Large outlays must be made for education, transportation, sources of power. and the like. Such investments require a relatively long time for pay-off, require large sums of money, and generally benefit the community as a whole. This indicates that government must generally play an important role in the process of providing money for social overhead. In fact, the most important precondition for take-off is often political. An effective government must maintain a tax and fiscal system which directs resources into modern uses and it is likely that only a vigorous central leadership can achieve this.

When the period of transition has begun, new types of enterprising men come forward and show an ability to raise money and a willingness to take risks in pursuit of profit or modernization. Banks appear, investment increases, and modern factories spring up. The people learn to operate a constantly changing economic system and come to accept progress as not only possible, but necessary. This activity may proceed, however, at a limited pace within a society mainly characterized by traditional, low productivity methods, and by the old social-political values and structures.

The Take-off

In the third stage, take-off, old resistance to steady growth is overcome and growth becomes the normal condition. Take-off is concentrated within two or three decades and its beginning can usually be traced to some sharp stimulus; for example, a political revolution, a technological improvement, a newly favorable international environment, or a shift to a very unfavorable position in terms of world trade. The most powerful single initiator of take-offs has been the railroad, which has performed the vital tasks of lowering internal transportation costs, developing a new export sector, and leading toward development of coal, iron, and engineering industries.

The following conditions are required for take-off: (a) a rise in the rate of productive investment to at least 10% of national income, (b) the development of one or more substantial manufacturing industries with a high rate of growth, and (c) the existence or quick emergence of a political, social, and institutional framework so developed as to keep up a continued growth. This further implies a capacity for raising money from domestic sources.

The take-off usually witnesses a social, political, and cultural victory for those who favor modernization of the economy over those who would either cling to the traditional society or seek other goals. New industries expand rapidly, encouraging still other industries, and increasing income in the hands of those who reinvest in the economy.

Drive to Technological Maturity

About forty years after a society ends take-off, technological maturity is usually achieved. During this

drive to maturity, the make-up of the economy changes constantly as technology improves. The economy finds it place in international trade. Goods formerly imported are produced at home. New import requirements develop along with new export commodities. Old industries level off and new industries accelerate, often with a shift toward more complex processes such as machine tools, chemicals, and electrical equipment. Thus, maturity is attained when an economy demonstrates its capacity to move beyond the original industries which powered its take-off and apply modern technology to virtually the whole range of its resources.

Three important non-economic aspects accompany the development of a maturing society.

First, the working force changes in composition, in real wages, in outlook, and in skills. By maturity, the percentage of the working force in agriculture has dwindled to a figure as low as 20% in many classes. Not only does the urban population grow, but also there is generally an increase in the proportion of white collar workers, highly trained technicians, and semi-skilled workers. These people realize that they can exert power, by organizing, to achieve higher real wages and greater security; hence, the process of moving toward maturity generates social and political pressures which lead toward humane modifications of the process.

Second, the character of the leader ship changes from the industrial tycoon to the efficient professional manager of a highly bureaucratized machine.

Third, the society as a whole takes for granted the miracle of industrialization and begins to question the merits of industrialization as an overriding objective. These changes pose new questions concerning future objectives.

High Mass-Consumption

In the final stage, the age of high mass-consumption, the society has ceased to accept the extension of modern technology as a primary objective. Real income per person increases and so does the effective demand for the products of a mature economy. Each society which has attained this stage of development has struck a unique balance, determined by geography, resources, values, and political leadership, among three broad objectives; (a) the

pursuit of external power; (b) the welfare state with a good deal of social legislation designed to redistribute income, to decrease working hours, and to increase social security in general; and (c) the expansion of consumer goods distribution.

Since growth normally proceeds by geometric progression, similar to a savings account if interest is left to compound with principal, the era of high mass-consumption will continue to gather momentum and vary its patterns.

Adapted from

W.W. Rostow, The Stages of Economic Growth. Cambridge Press, London, 1960, pp. 4-92.

PRECEDING PAGE BLANK-NOT FILMED

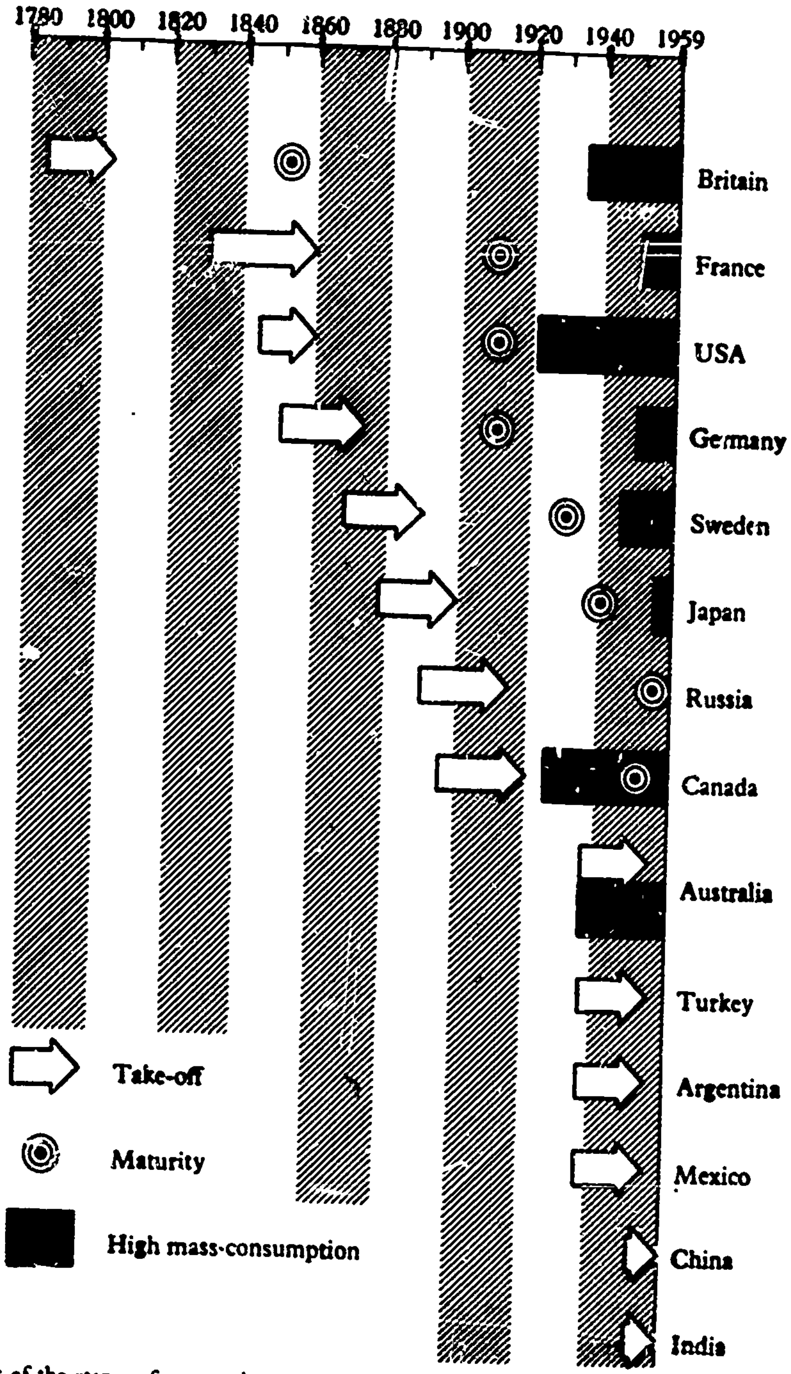


Chart of the stages of economic growth in selected countries. Note that Canada and Australia have entered the stage of high mass-consumption before reaching maturity. [By courtesy of the *Economist*.]

The length of each arrow indicates the length of the take-off stage.

1. The Stages of Economic Growth was written by
(K)
   1. Keynes
   2. Moore
   3. Rostow
   4. Stolurow

2. According to the article, how many stages of economic growth are there?
(K)
   1. 3
   2. 5
   3. 7
   4. 10

3. Which of the following is in the correct order?
(K)
   1. take-off, precondition for take-off, drive to maturity
   2. precondition for take-off, take-off, traditional
   3. take-off, drive to maturity, high mass-consumption
   4. precondition for take-off, high mass-consumption, drive to maturity

4. Rostow considers economic change to be the result of
(K)
   1. economic and social forces.
   2. economic and political forces.
   3. economic, political and social forces.
   4. "cold calculation" alone.

5. Approximately what per cent of the labor force of a traditional society does work related to agriculture?
(K)
   1. 20
   2. 35
   3. 50
   4. 75

6. Which stage of development is also known as the traditional period?
(K)
   1. traditional
   2. precondition for take-off
   3. take-off
   4. drive to technological maturity

7. The stage of development which frequently begins as a result of aggression is
(K)
   1. precondition for take-off
   2. take-off
   3. drive to technological maturity
   4. age of high mass-consumption

8. Which stage of growth is a nation's rate of investment under 5% of its national income?
(K) 1. traditional
2. take-off
3. drive to technological maturity
4. age of high mass-consumption

9. Which of the following is most likely to produce quick-yielding changes in productivity?
(K) 1. building a hydroelectric plant
2. manufacturing raw steel products
3. manufacturing heavy equipment
4. processing natural resources

10. Which of the following is a social overhead expense?
(K) 1. interstate highway
2. department store
3. watch factory
4. farm

11. The first stage in which growth becomes the normal condition is
(K) 1. precondition for take-off
2. take-off
3. drive to maturity
4. age of high mass-consumption

12. About how many decades does the take-off stage usually last?
(K) 1. one
2. two or three
3. four or five
4. between five and ten

13. What has been the most powerful single starter of the take-off stage?
(K) 1. railroads
2. agriculture
3. complex industries
4. coal, iron, and engineering industries

14. The minimum rate of productive investment of a country in the take-off stage is
(K) 1. 5%
2. 8%
3. 10%
4. 20%

15. Productive investment is measured as a percentage of
(K) 1. industrial earnings.
2. government earnings
3. national income
4. farm income

16. How many years does it usually take a nation to achieve maturity after it enters drive-to-maturity?
(K) 1. 20
2. 30
3. 40
4. 60

17. Economic growth usually progresses
(K) 1. geometrically.
2. arithmetically.
3. logarithmically.
4. inversely.

18. According to the chart, in 1959 India was in what stage?
(K) 1. traditional
2. precondition for take-off
3. take-off
4. maturity

19. According to the chart, in 1850 Britain was in what stage?
(K) 1. traditional
2. take-off
3. maturity
4. high mass-consumption

20. According to the chart, in which stage was Canada during the period 1900-1910?
(K) 1. traditional
2. take-off
3. maturity
4. high mass-consumption

21. Which of the following would most likely be a major export from a country in the transitional stage of economic growth?
(C) 1. farm machinery
2. paper products
3. furniture
4. crude oil

22. One trend appearing in the chart is that countries today, compared to countries a century ago
(C) 1. require more time for take-off.
2. seldom reach high mass-consumption.
3. move through the transitional stage more quickly.
4. move from take-off to high mass-consumption more quickly.

23. Which of the following occupations would most teenagers living in a traditional society desire?
(An) 1. factory manager
2. farm owner
3. bank president
4. research laboratory director

24. The statement, "two countries which violated the order of Rostow's stages of economic growth were Canada and Australia" is
(Ap) 1. true, according to the chart.
2. probably true, but the chart does not indicate it definitely.
3. false, according to the chart.
4. probably false, but the chart does not indicate it definitely.

25. According to the chart, in 1940 Sweden was in which stage of economic growth?
(C) 1. take-off
2. drive to maturity
3. maturity
4. high mass-consumption

26. The steam engine has been said to be the necessary foundation of the Industrial Revolution in Britain. According to the information given in the chart, the steam engine must have been invented
(An) 1. before 1785.
2. between 1786 and 1805.
3. between 1806 and 1855.
4. between 1856 and 1875.

27. According to the chart, a country can enter the stage of high mass-consumption
(C) 1. only after it reaches technological maturity.
2. without reaching technological maturity.
3. only after take-off has been completed.
4. without going through the traditional stage.

28. According to the chart, which of the following countries has probably been a major exporter of manufactured goods over the longest period of time?
(Ap) 1. Canada
2. France
3. Japan
4. Sweden

29. In which stage of economic development would the productive rate of investment most likely exceed 15% of the national income?
(C) 1. traditional
2. precondition for take-off
3. take-off
4. drive to maturity

* * * * * * * * *

NOTE: The following information must be used in answering questions 30-31: Suppose a country exists which has the following four characteristics:
a. It is very large and is controlled by a strong dictator.
b. 80% of the population are farmers.
c. The rate of investment is about 5% of the national income.
d. The education level of the general population is low.

30. This country is like countries in the traditional stage in how many of the four characteristics?
(Ap) 1. 1
2. 2
3. 3
4. 4

31. This country is like countries in high mass-consumption in how many of the four characteristics?
(Ap) 1. 0
2. 1
3. 2
4. 3

* * * * * * * * *

32. Which one of the following countries most likely has more than 50% of its labor force in agriculture at the present time?
(Ap) 1. Argentina
2. Australia
3. Canada
4. Sweden

33. A country which just built its first public university is probably in what stage?
(An) 1. precondition for take-off
2. take-off
3. drive to maturity
4. high mass-consumption

34. Suppose that in a particular country the "Jones Textile Company" has established three branch factories. This country is most likely in
(An) 1. precondition for take-off.
2. take-off.
3. drive to maturity.
4. high mass-consumption.

35. Mass hoarding of money during high mass-consumption would probably cause
(An) 1. increased production of consumer goods.
2. enactment of laws to provide higher unemployment benefits.
3. a decrease in the number of persons unemployed.
4. a reduction in the number of government employees.

36. Which of the following occupations would be most important in moving a country into the transitional stage?
(An) 1. lawyer
2. school teacher
3. mining engineer
4. electronics technician

37. In which stage of economic development is the relative power of the central government the weakest?
(C) 1. traditional
2. precondition for take-off
3. take-off
4. technological maturity

38. Suppose that in a particular country a 40-year bond plan has just been approved for obtaining money to build a hydroelectric plant. This country is most likely in
(An) 1. precondition for take-off.
2. take-off.
3. drive to maturity.
4. high mass-consumption.

39. Which of the following is not characteristic of a country in the traditional stage?
(Ap) 1. Average family size increased over last year.
2. 95% of the land is owned by 15% of the population.
3. National corn harvest was the largest ever.
4. Manufacture of tractors showed a 200% increase over last year.

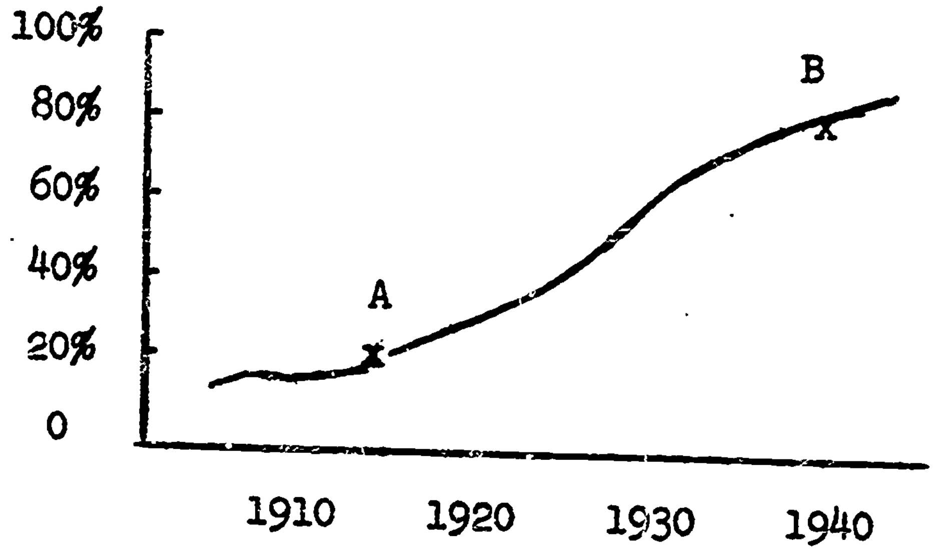
40. In the second paragraph of the passage Keynes said ".... as a result of cold calculation." What did he mean?
(C) 1. As a result of a desire to improve society.
2. As a result of objective planning.
3. As a result of regulatory legislation.
4. As a result of power seeking.

41. The statement, "the last country to reach the age of high mass-consumption was the United States" is
(Ap) 1. true, according to the chart.
2. probably true, but the chart does not indicate it definitely.
3. false, according to the chart.
4. probably false, but the chart does not indicate it definitely.

42. According to the chart, which one of the following countries probably has the highest per capita income?
(C) 1. Australia
2. China
3. Mexico
4. Canada

* * * * * * * *

NOTE: This graph, showing the per cent of non-agricultural workers in the total labor force each year in a particular country, must be used in answering questions 43-44.



43. According to the graph, which stage of growth is represented by point A?
(Ap)
1. traditional
2. precondition for take-off
3. take-off
4. drive to maturity

44. According to the graph, which stage of growth is represented by point B?

(Ap)
1. traditional
2. precondition for take-off
3. take-off
4. drive to maturity

* * * * * * * *

45. The statement, "of all the countries shown on the chart, Britain has the longest interval of time between maturity and high mass-consumption" is
(Ap)
1. true, according to the chart.
2. probably true, but the chart does not indicate it definitely.
3. false, according to the chart.
4. probably false, but the chart does not indicate it definitely.

46. Which of the following actions would do least to move a nation from the traditional to the transitional stage?
(Ap) 1. Strengthen the central government.
2. Encourage landowners to invest most of their income in industries.
3. Introduce modern farming techniques to increase production.
4. Reduce the compulsory retirement age.

47. Which of the following actions would most quickly move a country from the traditional to the transitional stage?
(C) 1. Build a technical college.
2. Have a war with a neighboring country that is in the maturity stage.
3. Hire an outstanding scientist from a country in the transitional stage.
4. Build a national system of airports.

48. Suppose that in a particular country labor unions representing clerks and secretaries have their maximum membership. This country is most likely in
(An) 1. precondition for take-off.
2. take-off.
3. drive to maturity.
4. high mass-consumption.

49. Which of the following countries reached maturity first?
(C) 1. Australia
2. Canada
3. France
4. Sweden

50. Unionization, the labor movement, or the banding together of workers to protect their rights is probably most characteristic of which stage?
(C) 1. traditional
2. take-off
3. drive to maturity
4. high-mass-consumption

51. Which of the following countries had the longest period of take-off?
(C) 1. Britain
2. Canada
3. France
4. Turkey

52. The statement, "there will be a longer period of time between the take-off and maturity for China than was required for Britain" is
(Ap)
1. true, according to the chart.
2. probably true, but the chart does not indicate it definitely.
3. false, according to the chart.
4. probably false, but the chart does not indicate it definitely.

53. The statement, "if Britain had started its take-off 60 years later, the time between its maturity and the age of high mass-consumption would have been shorter" is
(Ap)
1. true, according to the chart.
2. probably true, but the chart does not indicate it definitely.
3. false, according to the chart.
4. probably false, but the chart does not indicate it definitely.

54. Suppose you know the following information about a country:
a. half of the employed citizens are farmers;
b. the number of banks has doubled in the past 5 years.
c. its chief export is wheat and its chief import is farm machinery.

This country is probably in the
(Ap)
1. traditional stage
2. precondition for take-off stage.
3. take-off stage.
4. drive to maturity stage.

55. During drive-to-maturity, labor unions would
(C)
1. lose membership.
2. be controlled by the government.
3. change the social structure.
4. increase their influence.

56. Which of the following events would most likely move a country from the traditional to the transitional stage?
(Ap)
1. The discovery of a new strain of corn.
2. The overthrow of the government by army officers.
3. The establishment of a new industry to manufacture rifles.
4. The building of canals to provide inexpensive transportation.

57. The statement, "Australia will completely skip the maturity stage" is
(Ap)
1. true, according to the chart.
2. probably true, but the chart does not indicate it definitely.
3. false, according to the chart.
4. probably false, but the chart does not indicate it definitely.

58. In which stage is the economy most dependent on weather conditions?
(C)
1. traditional
2. precondition for take-off
3. take-off
4. drive to maturity

59. Suppose that in a particular country the government subsidizes the printing of books. This country is most likely in
(An)
1. precondition for take-off.
2. take-off.
3. drive to maturity.
4. high mass-consumption.

60. A country which had just won independence made up a constitution that included the following statements:
a. there will be no compulsory social security;
b. there will be no payments to unemployed persons;
c. the federal government will adopt a "hands off" policy toward the national economy.

Which of the following statements would best describe the economic characteristics of the country?
1. There are no large land owners.
2. Most of the population is employed in industry.
3. Over 50% of the population are involved in agriculture.
4. The rate of investment is 15% of the national income.

61. A country that has 40% of its population working in agriculture and the remainder working in heavy industries (machines, locomotives, etc.) is ruled by a dictator. The country has been in maturity for 60 years. Which of the following courses of action would probably lead to movement into high mass-consumption?
(An) 1. Overthrow the dictator.
2. Start a war with a neighboring country.
3. Manufacture and sell consumer goods.
4. Make the agricultural program more efficient.

62. In which stage would the occupation of "social worker" be the most common?
(Ap) 1. transitional
2. take-off
3. drive to maturity
4. high mass-consumption

63. The movement of France and Germany from maturity to high mass-consumption might have been delayed by the
(An) 1. Russian Revolution.
2. Second World War.
3. lack of scientific developments useful in industrialization.
4. high proportion of citizens engaged in agriculture.

64. Rostow states that when scientific methods are applied to agriculture, the proportion of people needed in agriculture decreases. He leads the reader to believe that the displaced agricultural workers will be
(An) 1. more or less permanently unemployed.
2. given jobs by the government.
3. absorbed into industrial activities.
4. employed in agricultural activities by less economically advanced countries.

65. Which of the following is <u>least</u> characteristic of a country in the stage of maturity?
(C) 1. Half of the population is involved in agriculture.
2. Demands for consumer goods exceed production.
3. Factories are expanding rapidly.
4. Heavy industry is emphasized.

66. Which of the following inventions probably had the most profound effect on the economic activities of man?
(An) 1. water wheel
2. steam engine
3. telephone
4. airplane

67. Strong labor unions will be found only in countries in which stage(s) of development?
(An) 1. high mass-consumption
2. drive to maturity
3. drive to maturity and high mass-consumption
4. take-off, drive to maturity, and high mass-consumption

68. When the United States gives economic assistance to an underdeveloped country, that aid is least likely intended to change the country from
(An) 1. traditional to transitional.
2. transitional to take-off.
3. take-off to maturity.
4. maturity to high mass-consumption.

69. Take-off, according to Rostow, is brought about by a sharp stimulus such as those listed below. Which one probably started take-off in the United States?
(An) 1. political revolution
2. technological improvement
3. newly favorable international environment
4. shift to unfavorable world trade conditions

70. Before take off can begin the citizens must accept
(C) 1. a lower standard of living.
2. government control of business.
3. control by the land owners.
4. constant change.

71. Which of the following is the most probable reason take-off ended in the United States when it did?
(An) 1. the Civil War started
2. the Gold Rush started
3. the east-west railroad was completed
4. the westward expansion was just getting under way

72. Based on the information in the chart, probably the process of industrialization is least complete in
(C) 1. Australia
2. Canada
3. France
4. Sweden

73. If progress is defined as movement from the traditional society to the age of high mass-consumption, then progress does not necessarily provide for increased
(C) 1. leisure time.
2. material goods.
3. religious dedication.
4. standard of living.

74. Several countries, for example, the United States and France, have foreign aid programs which are intended to stimulate the technical and industrial progress of underdeveloped nations. Which one of the following reasons for their doing so agrees best with the information in the reading passage.
(An) 1. Wealthy countries are morally obligated to help poor countries.
2. New markets and trade agreements can be developed in the under-developed countries.
3. World power can be gained by installing your political viewpoints in the under-developed countries.
4. Foreign aid is a convenient way of getting rid of surplus products.

75. According to the chart, the Canadian nation-wide railroad system was probably completed about
(C) 1. 1820
2. 1840
3. 1900
4. 1920

76. The statement, "Turkey will never enter the stage of high mass-consumption" is
(Ap) 1. true, according to the chart.
2. probably true, but the chart does not indicate it definitely.
3. false, according to the chart.
4. probably false, but the chart does not indicate it definitely.

77. In which stage of economic development would the occupation of "college professor" be least common?
(Ap) 1. traditional
2. precondition for take-off
3. take-off
4. drive to maturity

78. One characteristic of the high mass-consumption stage is increased social legislation to redistribute income. Which one of the following measures would probably be least useful in redistributing income?
(An) 1. An annual tax that is based on income.
2. A 3% sales tax on all goods and services.
3. Regular payments of money to the unemployed.
4. An annual tax that is the same for all residents.

79. Which one of the following countries did not pass through the stages of economic growth in the usual order?
(C) 1. Australia
2. Britain
3. India
4. United States

80. Prior to the Civil War, cotton was the most important product of the South. What is the most advanced stage in which the pre-Civil War South could be properly classified?
(Ap) 1. traditional
2. precondition for take-off
3. take-off
4. drive to maturity

STAGES OF ECONOMIC GROWTH
(Booklet Number 2)

Directions

Your answers to each of items 81-95 should be written directly on the test booklet in the space provided below each item. Should you need additional space for an item, write "continued"; then write the remainder of your answer on the back of the last page of the booklet.

You received with this test booklet another booklet which contains the article about Stages of Economic Growth. You may refer to it any time during the examination. You might wish to scan the article after you read these directions to refresh your memory.

Your answers to these items will depend on your understanding and thoughts about the article. Each of you might write an entirely different answer to an item because there are many possible answers to each of them. So please write whatever you believe to be appropriate.

Answers to these items do not appear as such in the article. Therefore, you will certainly answer incorrectly if you copy material directly from the article.

81. Suppose that you are a national leader in Australia.
(Sy) Describe some of the things that could be done to enable the country to reach maturity.

82. Political campaigns are based on issues that are
(Sy) important to the citizens. If you were running for a national office in a country which was in maturity what promises would you make in an effort to win the election?

83. The lists of manufactured goods of countries in take-
(Sy) off and in high mass-consumption are similar in many ways and are different in many ways. Make two short lists of manufactured goods, one representative of the manufactured goods of a country in high mass-consumption. How are your two lists alike and how are they different?

84. According to the chart, not all countries spent the
(Sy) same amount of time in each stage, nor did they progress through the stages in the same order. Briefly describe conditions under which a country might be able to skip or combine stages.

85. (Sy) Suppose that you are president of a country which is poor, underdeveloped, and in the transitional stage. A wealthy nation has offered aid in any one of the following forms: machines, professional and technical personnel, surplus food, or household goods and clothing. Choose the one form of aid which you think would produce the greatest long term benefits for your country, and tell why that form would be more beneficial than the others.

86.-95. Suppose that you are in doubt about some aspects of Rostow's theory. Listed below are several statements which could be true. Assume each statement to be true. Do not consider any relationships which might exist between the statements.

On the line following the statement place an X in the space after:

"greater trust"--if the statement leads you to have greater trust in Rostow's theory.

"less trust"--if the statement leads you to have less trust in Rostow's theory.

"no effect"--if the statement has no bearing on your trust in Rostow's theory.

In the space provided below each item write a brief statement giving the reason(s) why you marked the answer you did.

86. (Ev) In the mid-1930's Japan was better prepared for war than was China.
greater trust ___ less trust ___ no effect ___

87. (Ev) Rostow is an economic advisor to the President of the United States.
greater trust ___ less trust ___ no effect ___

88 (Ev) Several great civilizations have developed and disappeared in the course of world history.
greater trust ___ less trust ___ no effect ___

89. (Ev) A particular state in the United States has 80% of its population engaged in agriculture; however, the wealth and state governmental control are in the hands of several dozen large land owners.
greater trust ___ less trust ___ no effect ___

90. (Ev) There are several poorly developed countries that have been completely controlled by strong dictators for several decades.
greater trust ___ less trust ___ no effect ___

91. (Ev) Mr. Rostow is a college student who wrote the article as an assignment for a course in business economics.
greater trust ___ less trust ___ no effect ___

92. (Ev) Several new nations are being created in Africa.
greater trust ___ less trust ___ no effect ___

93. (Ev) Mr. Rostow has two college degrees.
greater trust ___ less trust ___ no effect ___

94. (Ev) Mr. Rostow wrote his theory and drew the chart in the reading passage in 1930.
greater trust ___ less trust ___ no effect ___

95. (Ev) The United States has a surplus of agricultural products.
greater trust ___ less trust ___ no effect ___

TABLE EG-1

STAGES OF ECONOMIC GROWTH

Percentiles*

Raw Score	Knowledge 9	Knowledge 10	Knowledge 11	Knowledge 12	Comprehension 9	Comprehension 10	Comprehension 11	Comprehension 12
20	95	93	91	88				
19	84	78	74	71			99	98
18	72	63	58	54		99	97	96
17	61	50	45	40	99	96	92	93
16	52	41	35	31	97	93	88	88
15	43	34	27	24	95	89	82	81
14	36	29	21	18	92	84	76	75
13	30	24	16	15	88	77	69	66
12	26	22	13	11	81	69	62	58
11	21	17	10	09	74	62	54	49
10	18	14	07	07	66	52	45	38
9	12	10	06	05	57	43	36	31
8	08	07	04	04	46	35	30	23
7	05	04	03		35	26	22	16
6	03	03	02	02	24	17	15	10
5	02	02	01	01	14	11	09	06
4	01				08	05	05	03
3					03	02	02	
2					01	01	01	01
1								
0								

*Percentiles computed on the per cent of students scoring below a given raw score. Percentiles for this form were based on the following numbers of students:

Grade 9 - 1199
Grade 10 - 1451
Grade 11 - 1283
Grade 12 - 1194

TABLE EG-2

STAGES OF ECONOMIC GROWTH

Percentiles

Raw Score	Application 9	Application 10	Application 11	Application 12	Analysis 9	Analysis 10	Analysis 11	Analysis 12
20								
19								
18			99	99				
17	99	98	96	96				
16	98	96	92	91				
15	96	93	85	84		99	98	98
14	92	86	78	76		98	95	94
13	87	80	71	67	98	95	91	88
12	82	70	62	57	96	91	85	81
11	75	62	54	46	93	85	78	73
10	67	53	44	38	88	77	69	64
9	57	43	37	31	79	67	59	55
8	47	35	28	24	69	57	49	45
7	36	26	20	17	57	46	39	34
6	24	18	14	11	42	35	28	24
5	15	10	09	07	28	22	20	15
4	08	06	06	03	17	12	11	09
3	04	02	02	01	09	06	05	04
2	01	01	01		03	02	03	02
1					01			
0								

TABLE EG-3

STAGES OF ECONOMIC GROWTH

Percentiles

Raw Score	Synthesis 9	10	11	12	Evaluation 9	10	11	12
20								
19								
18								
17								
16								
15								
14								
13								
12								
11								
10								
9								
8			99	99				
7		99	98	98				
6		98	95	97				99
5	99	96	92	93		99	99	98
4	97	92	86	87	99	98	97	96
3	94	87	76	77	97	94	93	91
2	86	76	62	61	92	89	85	80
1	69	57	40	37	81	77	69	62
0	36	32	18	15	59	55	44	38

PRECEDING PAGE BLANK-NOT FILMED

TABLE EG-4

Item Data

Stages of Economic Growth
Knowledge

Item No.	Key	Difficulty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
1	3	.96	.25	.25	.03	.01	.96	.01	.00
2	2	.83	.39	.39	.10	.83	.02	.01	.04
3	3	.76	.43	.43	.05	.10	.76	.08	.01
4	3	.77	.49	.49	.09	.06	.77	.08	.00
5	4	.84	.55	.54	.09	.04	.02	.84	.01
6	2	.80	.58	.57	.09	.80	.06	.05	.00
7	1	.67	.59	.58	.67	.15	.10	.07	.00
8	1	.73	.51	.51	.73	.17	.05	.04	.01
9	4	.63	.36	.36	.15	.10	.10	.63	.01
10	1	.69	.45	.44	.69	.10	.07	.14	.01
11	2	.66	.57	.57	.12	.66	.15	.07	.00
12	2	.85	.55	.54	.06	.85	.06	.03	.00
13	1	.78	.61	.61	.78	.08	.06	.08	.00
14	3	.83	.61	.60	.09	.03	.83	.05	.00
15	3	.63	.40	.40	.23	.07	.63	.07	.01
16	3	.78	.61	.60	.10	.06	.78	.05	.00
17	1	.64	.47	.47	.64	.11	.11	.12	.01
18	3	.83	.44	.44	.04	.10	.83	.03	.00
19	3	.82	.42	.42	.03	.07	.82	.09	.00
20	2	.87	.46	.46	.04	.87	.05	.04	.00

Raw Scores		Corrected Scores	
Mean	15.34	Mean =	13.821
S.D. =	3.96	S.D. =	5.28
KR #20 =	0.824	KR #20 =	0.924

N = 5143

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

TABLE EG-5

Item Data

Stages of Ecnomic Growth
Comprehension

Item No.	Key	Diffi- culty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
21	4	.37	.36	.37	.37	.16	.09	.37	.01
22	4	.50	.37	.37	.18	.12	.20	.50	.01
25	3	.63	.42	.43	.05	.15	.63	.17	.01
27	2	.40	.42	.44	.19	.40	.25	.15	.01
29	4	.52	.40	.41	.07	.14	.27	.52	.01
37	1	.47	.37	.39	.47	.22	.16	.14	.01
40	2	.44	.32	.32	.24	.44	.11	.20	.01
42	4	.74	.41	.41	.12	.08	.05	.74	.01
47	2	.24	.23	.23	.40	.24	.14	.19	.02
49	3	.62	.44	.43	.08	.22	.62	.05	.03
50	3	.41	.34	.34	.11	.23	.41	.22	.03
51	3	.61	.49	.49	.22	.07	.61	.08	.03
55	4	.58	.52	.51	.09	.12	.17	.58	.04
58	1	.57	.57	.57	.57	.19	.13	.07	.05
65	1	.54	.57	.55	.54	.18	.10	.09	.09
70	4	.52	.55	.51	.09	.16	.12	.52	.11
72	1	.31	.33	.30	.31	.10	.30	.16	.13
73	3	.31	.42	.39	.24	.14	.31	.18	.14
75	3	.27	.25	.22	.07	.13	.27	.38	.15
79	1	.55	.54	.49	.55	.10	.12	.06	.17

Raw Scores

Mean = 9.61
S.D. = 3.99
KR #20 = 0.745

Corrected Scores

Mean = 6.48
S.D. = 5.14
KR #20 = 0.867

N = 5143

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

TABLE EG-6

Item Data

Stages of Economic Growth
Application

Item No.	Key	Difficulty	r_1*	r_2*	Proportions Choosing Each Response 1	2	3	4	Omit
24	1	.68	.50	.51	.68	.14	.11	.06	.01
28	2	.43	.34	.35	.44	.43	.09	.05	.00
30	3	.40	.28	.28	.12	.31	.40	.16	.01
31	1	.61	.45	.46	.61	.20	.12	.07	.01
32	1	.61	.54	.55	.61	.16	.16	.07	.01
39	4	.45	.43	.44	.15	.27	.12	.45	.01
41	3	.79	.51	.50	.07	.07	.79	.06	.01
43	1	.36	.43	.43	.36	.27	.23	.13	.01
44	4	.69	.43	.42	.13	.06	.10	.69	.01
45	1	.74	.47	.46	.74	.11	.09	.04	.02
46	4	.52	.46	.46	.15	.13	.18	.52	.02
52	2	.39	.30	.30	.19	.39	.15	.23	.03
53	2	.43	.38	.37	.22	.43	.13	.18	.04
54	2	.39	.39	.39	.13	.39	.23	.21	.04
56	4	.41	.33	.31	.17	.19	.18	.41	.05
57	4	.16	.19	.18	.48	.21	.10	.16	.05
62	4	.29	.38	.37	.18	.19	.27	.29	.07
76	4	.36	.51	.48	.15	.24	.09	.36	.15
77	1	.51	.52	.48	.51	.14	.10	.09	.16
80	2	.27	.29	.26	.27	.27	.18	.10	.17

Raw Scores		Corrected Scores	
Mean	= 9.47	Mean	= 6.24
S.D.	= 3.82	S.D.	= 4.96
KR #20	= 0.731	KR #20	= 0.862

N = 5143

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

TABLE EG-7

Item Data

Stages of Economic Growth
Analysis

Item No.	Key	Difficulty	r_1*	r_2*	Proportion Choosing Each Response 1	2	3	4	Omit
23	2	.58	.38	.42	.11	.58	.19	.12	.00
26	1	.27	.24	.26	.27	.26	.31	.15	.01
33	2	.26	.18	.18	.40	.26	.28	.06	.01
34	3	.35	.18	.18	.10	.27	.35	.27	.01
35	2	.29	.23	.25	.28	.29	.26	.16	.01
36	3	.32	.20	.21	.10	.37	.32	.20	.01
38	1	.21	.15	.16	.21	.23	.37	.17	.01
48	4	.39	.35	.36	.11	.16	.31	.39	.02
59	4	.14	.08	.06	.36	.24	.21	.14	.05
60	3	.41	.40	.39	.12	.26	.41	.15	.06
61	3	.30	.27	.24	.35	.12	.30	.16	.07
63	2	.50	.41	.37	.08	.50	.21	.13	.07
64	3	.47	.52	.49	.18	.14	.47	.13	.08
66	2	.53	.51	.47	.12	.53	.12	.15	.09
67	3	.41	.46	.43	.17	.13	.41	.19	.09
68	4	.37	.53	.50	.18	.18	.17	.37	.10
69	2	.34	.38	.35	.22	.34	.20	.13	.11
71	1	.40	.47	.43	.40	.12	.23	.13	.12
74	2	.41	.43	.39	.10	.41	.23	.11	.15
78	4	.33	.50	.47	.13	.17	.20	.33	.17

Raw Scores

Mean = 7.28
S.D. = 3.26
KR #20 = 0.614

Corrected Scores

Mean = 3.44
S.D. = 4.01
KR #20 = 0.763

N = 5143

*r_1 = Point biserial correlation, raw scores
r_2 = Point biserial correlation, corrected scores

Appendix E

Administering the Tests

The preparations for administering the taxonomy-type tests are similar to those for any standardized test.

The tests should be administered to classroom-size groups of students rather than to massed groups of students as often occurs when tests are administered in school cafeterias, lunchrooms, and auditoriums. The small-group situation will provide the examiner with a maximum of control and the students with a minimum of distractions. The former is especially important because the taxonomy-type tests differ somewhat in format from typical standardized tests and the desired student responses are substantially different. The small-group situation will enable responding to students' questions during the administration with a minimum of distraction of others. Also, it will facilitate proctoring the examination, a part of which should be observing students to note whether they are complying with the directions.

Each test consists of two parts. Part A contains a reading passage and eighty multiple-choice items. Part B contains fifteen free-response items. Part A must be administered before Part B. A short break may be allowed between the Parts but it is not necessary to do so.

The total time for administering each test (Parts A and B) is about two hours. Experience has shown that sixty minutes are sufficient for most students to complete Part A. The range of completion times for Part B is forty to sixty minutes. A time limit of fifty minutes should be announced for Part B. If most students finish before then, time can be called. If few are finished at the end of fifty minutes, then the time limit can be extended for ten more minutes.

Because the content of the reading passages will be unfamiliar to students, they should be instructed to read the entire passage before attempting to respond to the items. The administrator and/or proctors should enforce this instruction! It is recommended that the administrator remind students, while they are taking Part A, that it is permissible to refer to the reading passage at any time. Reference to the reading passage at any time is also permitted during the administration of Part B.

The following materials will be required for each pupil for each test: (1) Part A booklet (2) Separate answer sheet if responses are not made directly on booklet.

(3) Several pencils and a special pencil if required for the special answer sheet (4) Part B booklet.

The administrator should have a copy of the Directions for administering the test and copies of Parts A and B of the test.

The administrator should have some reliable method of keeping time, but a stopwatch or interval timer is unnecessary.

The following pages contain the set of directions which was used for the experimental administration of the tests during the project. They are presenced as they were used and as a guide for anyone who might wish to use the tests. They may be modified to fit the situation in which the tests will be used.

INSTRUCTIONS FOR ADMINISTERING THE TESTS

First Administration of any Test

Read: Today you are going to take some tests of your ability to apply skills and understandings which you have been learning since you first entered school. These tests were developed at Florida State University and you are helping to establish how well the average high school student can do on them. Because these tests are being used for research, it is very important that you do your very best on them.

When taking each test you will be asked to read an article and then to answer questions about it. These tests are not like most tests you have taken because they are not based on specific subjects you have studied.

We will now distribute the materials for the test. Please leave them on your desks until you are given further instruction.

(Allow time for distribution of Part A and answer sheets)

Read: You must use a number two pencil. If you do not have a number two pencil, raise your hand.

(Pause)

Read: Now look at your answer sheet. Fill in the following spaces: name, grade, school, and name of test. You will find the name of the test at the top of the cover page of the test booklet (at this point, give the names of the tests being administered in the session).

Read: We will now read the instructions on the front cover of the test booklet. Follow the instructions as I read them aloud:

This is a test of your ability to apply skills and understandings which you have been learning since you entered school. You will be asked to read a passage called ____________ and then you will be asked to answer questions about it. You may refer to the passage at any time during the test.

Your answers must be marked on the answer sheet which has been provided for you. Use only a number 2 pencil. (Your booklet says to use only the pencil provided, but any number 2 pencil is okay. Do not use any other number pencil, do not use ballpoint or fountain pens.) Do not make any stray marks. If you make an error, erase completely before marking your new answer.

Please make no marks on the test booklet.

The following is a sample question to show you how your answers are to be marked. Study the sample and if you have any questions raise your hand.

(Pause while students study sample).

Read: You should answer as many questions as you can. Do not spend a great amount of time on any one question. If you are in doubt about the answer to a question, then guess. Remember that you may refer to the reading passage at any time during the test.

When taking this test you will notice that the answers to some of the questions appear directly in the article. The answers to other questions will require some thinking on your part. When you have finished reading the passage begin answering the questions. Turn the page and start reading. (Students will have 60 minutes to read the passage and answer the questions.)

At the end of 60 minutes say:

Read: "Stop." We will now collect the answer sheets. Do not turn in your test booklet as you will need it for the next part of this test.

(Collect answer sheets and distribute Part B, do not collect Part A)

Read: Place your name, identification number, school, grade, and date in the spaces provided on the booklet. We will now read the instructions for this test.

Your answers to each of items 81-95 should be written directly below each item. Should you need additional space for an item write "continued" then write the remainder of your answer on the back of the last page of the booklet.

You received with this test booklet another booklet which contains the article about ________. You may refer to it at any time during the examination. You might wish to scan the article after you read these directions to refresh your memory.

Your answers to these items will depend on your understanding and thoughts about the article. Each of you might write an entirely different answer to an item because there are many possible answers to each of them. So please write whatever you believe to be appropriate.

Answers to these items do not appear as such in the article. Therefore, you will certainly answer incorrectly if you copy the material directly from the article.

Note to Administrator: The range of completion times for Part B is forty to sixty minutes. A time limit of fifty minutes should be announced for Part B. If most students finish before then, time can be called. If few finish at the end of fifty minutes then the time limit can be extended for ten more minutes.

At the end of 50 minutes say: Stop. We will now collect the test you have just completed.

(Collect Part A and Part B separately)

INSTRUCTIONS FOR ADMINISTERING THE TESTS

Second, Third, and Fourth Administrations

Read: This test is similar to the other one(s) you have taken in this series. You will again be asked to read an article and answer questions about it.

Read: We will now distribute the materials for the test. Please leave them on your desks until you are given further instruction. (Allow time for distribution of Part A and answer sheets)

Read: You must use a number two pencil. If you do not have a number two pencil raise your hand. (Repeat the appropriate directions from the previous page.)

Scoring the Tests

Two types of student responses are elicited by each taxonomy-test. Part A of each is multiple-choice and Part B is free-response.

Part A of each form requires four scoring keys-- Knowledge, Comprehension, Application, and Analysis. Scoring keys can be constructed readily from the following tables which contain for each form, the item number, the keyed response, and the taxonomy level. For scoring small numbers of answer sheets, a punch matrix overlay and hand-scoring will be feasible. For large numbers of answer sheets, IBM 1230 or Digitex special scoring keys, or computer programs will be more feasible and economical than handscoring.

The items of Part B elicit written responses from students. These must be scored by judges who are thoroughly familiar with the processes of Synthesis and Evaluation, the reading passages, and the pertinent items.

It is recommended that the scoring be done by teams of three or four judges each. They should first become knowledgeable about the tests and processes described above, then they should become equally familiar with the scoring directions and scored sample responses which appear in this appendix. Then, a sample of responses should be drawn from those they will score and these should be scored independently by each of them. Inter-judge reliabilities should be calculated. If the magnitude of it is unsatisfactory, then additional training should be give to them. When a satisfactory level of agreement is attained, then scoring should begin. It is recommended that quality control studies be done periodically during scoring to assure that this level is maintained.

The scored responses which are included here for each quality level of each item of the four forms should be helpful to judges during training and when they are actually rating responses. However, it should be noted that these are merely a sample of the responses which students actually gave to the items. Undoubtedly, different responses will be commonplace. Therefore, the principles underlying the scoring procedures will be much more valuable than specific knowledge of how each sample item was scored.

In Section I, following, the necessary information for scoring Part A of each form is provided. Section II contains the definitions of Synthesis and Evaluation along with general characteristics which determine the score for each item. Sections III through VI contain scored sample responses for Synthesis and Evaluation items for each of the four forms.

SECTION I

Key

Atomic Structure

Part A

Item Number	Correct Response	Process Level	Item Number	Correct Response	Process Level
1	1	Knowledge	41	1	Application
2	1	Knowledge	42	4	Comprehension
3	2	Knowledge	43	3	Application
4	3	Knowledge	44	1	Application
5	2	Knowledge	45	4	Application
6	2	Knowledge	46	3	Application
7	4	Knowledge	47	2	Comprehension
8	3	Knowledge	48	2	Analysis
9	4	Knowledge	49	1	Comprehension
10	2	Knowledge	50	3	Comprehension
11	4	Knowledge	51	1	Comprehension
12	2	Knowledge	52	2	Application
13	3	Knowledge	53	2	Application
14	1	Knowledge	54	2	Application
15	1	Knowledge	55	4	Comprehension
16	4	Knowledge	56	3	Application
17	3	Knowledge	57	1	Application
18	4	Knowledge	58	2	Comprehension
19	1	Knowledge	59	2	Analysis
20	2	Knowledge	60	3	Analysis
21	1	Comprehension	61	1	Analysis
22	3	Comprehension	62	1	Application
23	1	Analysis	63	4	Analysis
24	1	Application	64	2	Analysis
25	3	Comprehension	65	3	Comprehension
26	3	Analysis	66	1	Analysis
27	4	Comprehension	67	4	Analysis
28	1	Application	68	4	Analysis
29	4	Comprehension	69	2	Analysis
30	2	Application	70	2	Comprehension
31	4	Application	71	4	Analysis
32	3	Application	72	4	Comprehension
33	2	Analysis	73	3	Comprehension
34	4	Analysis	74	1	Analysis
35	1	Analysis	75	4	Comprehension
36	1	Analysis	76	3	Application
37	2	Comprehension	77	2	Application
38	4	Analysis	78	3	Analysis
39	1	Application	79	1	Comprehension
40	4	Comprehension	80	3	Application

Key

Glaciers

Part A

Item Number	Correct Response	Process Level	Item Number	Correct Response	Process Level
1	4	Knowledge	41	2	Application
2	3	Knowledge	42	2	Comprehension
3	1	Knowledge	43	1	Application
4	1	Knowledge	44	2	Application
5	3	Knowledge	45	2	Application
6	1	Knowledge	46	2	Application
7	3	Knowledge	47	1	Comprehension
8	4	Knowledge	48	3	Analysis
9	2	Knowledge	49	3	Comprehension
10	2	Knowledge	50	3	Comprehension
11	4	Knowledge	51	1	Comprehension
12	3	Knowledge	52	2	Application
13	4	Knowledge	53	2	Application
14	2	Knowledge	54	1	Application
15	2	Knowledge	55	2	Comprehension
16	1	Knowledge	56	1	Application
17	1	Knowledge	57	1	Application
18	4	Knowledge	58	2	Comprehension
19	1	Knowledge	59	3	Analysis
20	2	Knowledge	60	1	Analysis
21	4	Comprehension	61	4	Analysis
22	2	Comprehension	62	1	Application
23	3	Analysis	63	4	Analysis
24	2	Application	64	3	Analysis
25	4	Comprehension	65	3	Comprehension
26	1	Analysis	66	4	Analysis
27	4	Comprehension	67	1	Analysis
28	4	Application	68	2	Analysis
29	2	Comprehension	69	4	Analysis
30	4	Application	70	1	Comprehension
31	3	Application	71	2	Analysis
32	3	Application	72	3	Comprehension
33	4	Analysis	73	3	Comprehension
34	2	Analysis	74	4	Analysis
35	3	Analysis	75	2	Comprehension
36	1	Analysis	76	3	Application
37	1	Comprehension	77	1	Application
38	4	Analysis	78	2	Analysis
39	4	Application	79	3	Comprehension
40	3	Comprehension	80	4	Application

Key

Lisbon Earthquake

Part A

Item Number	Correct Response	Process Level	Item Number	Correct Response	Process Level
1	3	Knowledge	41	3	Application
2	4	Knowledge	42	2	Comprehension
3	2	Knowledge	43	2	Application
4	2	Knowledge	44	1	Application
5	1	Knowledge	45	4	Application
6	1	Knowledge	46	4	Application
7	1	Knowledge	47	1	Comprehension
8	3	Knowledge	48	1	Analysis
9	2	Knowledge	49	2	Comprehension
10	1	Knowledge	50	4	Comprehension
11	3	Knowledge	51	3	Comprehension
12	1	Knowledge	52	4	Application
13	2	Knowledge	53	3	Application
14	4	Knowledge	54	3	Application
15	2	Knowledge	55	2	Comprehension
16	3	Knowledge	56	2	Application
17	4	Knowledge	57	1	Application
18	2	Knowledge	58	3	Comprehension
19	3	Knowledge	59	4	Analysis
20	1	Knowledge	60	1	Analysis
21	4	Comprehension	61	3	Analysis
22	1	Comprehension	62	1	Application
23	4	Analysis	63	2	Analysis
24	3	Application	64	3	Analysis
25	4	Comprehension	65	3	Comprehension
26	3	Analysis	66	1	Analysis
27	3	Comprehension	67	4	Analysis
28	2	Application	68	4	Analysis
29	4	Comprehension	69	2	Analysis
30	1	Application	70	3	Comprehension
31	4	Application	71	3	Analysis
32	4	Application	72	1	Comprehension
33	2	Analysis	73	1	Comprehension
34	2	Analysis	74	4	Analysis
35	1	Analysis	75	3	Comprehension
36	1	Analysis	76	2	Application
37	1	Comprehension	77	4	Application
38	2	Analysis	78	4	Analysis
39	2	Application	79	2	Comprehension
40	4	Comprehension	80	1	Application

Key

Economic Growth

Part A

Item Number	Correct Response	Process Level	Item Number	Correct Response	Process Level
1	3	Knowledge	41	3	Application
2	2	Knowledge	42	4	Comprehension
3	3	Knowledge	43	1	Application
4	3	Knowledge	44	4	Application
5	4	Knowledge	45	1	Application
6	2	Knowledge	46	4	Application
7	1	Knowledge	47	2	Comprehension
8	1	Knowledge	48	4	Analysis
9	4	Knowledge	49	3	Comprehension
10	1	Knowledge	50	3	Comprehension
11	2	Knowledge	51	3	Comprehension
12	2	Knowledge	52	2	Application
13	1	Knowledge	53	2	Application
14	3	Knowledge	54	2	Application
15	3	Knowledge	55	4	Comprehension
16	3	Knowledge	56	4	Application
17	1	Knowledge	57	4	Application
18	3	Knowledge	58	1	Comprehension
19	3	Knowledge	59	4	Analysis
20	2	Knowledge	60	3	Analysis
21	4	Comprehension	61	3	Analysis
22	4	Comprehension	62	4	Application
23	2	Analysis	63	2	Analysis
24	1	Application	64	3	Analysis
25	3	Comprehension	65	1	Comprehension
26	1	Analysis	66	2	Analysis
27	2	Comprehension	67	3	Analysis
28	2	Application	68	4	Analysis
29	4	Comprehension	69	2	Analysis
30	3	Application	70	4	Comprehension
31	1	Application	71	1	Analysis
32	1	Application	72	1	Comprehension
33	2	Analysis	73	3	Comprehension
34	3	Analysis	74	2	Analysis
35	2	Analysis	75	3	Comprehension
36	3	Analysis	76	4	Application
37	1	Comprehension	77	1	Application
38	1	Analysis	78	4	Analysis
39	4	Application	79	1	Comprehension
40	2	Comprehension	80	2	Application

SECTION II

Definition of Synthesis[1]

Synthesis is defined as putting together of elements and parts so as to form a whole. The task of the student is to combine certain elements to constitute a pattern of structure not clearly there before. It will generally be necessary for the student to combine parts of the reading passage with generalizations from their prior learning in order to form a well integrated product.

Essay questions, as used in achievement and aptitude tests, generally require only the expression of remembered ideas and interpretation of given materials. Synthesis responses require uniqueness and originality, requiring the student to draw upon elements from many sources to construct a new product instead of studying a simple whole and merely reproducing elements of that whole. Comprehension, Application and Analysis are, of course, involved in a synthesis product, but products stopping at any of these levels are partial products and are less complete than the desired product.

There are three different kinds of Synthesis, which are on the basis of the product required by the task. These are (1) production of a unique communication, (2) production of a plan or proposed set of operations, and (3) production of a set of abstract relations.

A unique communication, is a product through which the student is trying to communicate his ideas and experiences to the reader. The mode of expression is less conventional than the typical essay question requires and may be in the form of a cartoon or illustration, or a descriptive story or essay. The key is the effectiveness with which the student uses the particular mode of expression. The mode also sets limits within which the student must accomplish the purposes set by the task.

The second type of product is a plan or proposed set of operations. The product must satisfy the requirements of the task which are usually in the form of data or specifications to be taken into account by the student. The requirements of the task furnish a rather well-defined criterion against which the student's product may be evaluated. The product must not be viewed strictly by preconceived standards however, because a student may conceive of a plan or process that is uniquely his own.

[1]The material in this section is based on the discussion of Synthesis in Part II, Section 5 of the Taxonomy of Educational Objectives, Handbook I, Bloom, et al.

The third type of Synthesis product includes objectives that require the student to produce, or derive a set of abstract relations. The student may begin with concrete data which he proceeds to classify or explain, or the student may begin with some basic proposition and deduce other propositions or relations. In either case the product must fit the facts, adequately account for seemingly unrelated phenomena, and reason out the task in terms of given information. The task permits the student to carry his thinking quite far, but the specific information creates a rigorous objective criteria which his product of Synthesis must meet.

Response Characteristics

Synthesis items are scored on a scale ranging from 0 to 4. The general characteristics of the responses in each category are presented below.

Score 0 A response which shows no attempt to complete the task set forth in the item. This would include responses which consist entirely of copied material, of the problem posed, take issue with the "facts" of the item, or show erroneous interpretation of the task.

Score 1 A response which would indicate that the subject has a basic understanding of the task at hand. This would include responses which are mere listings of the elements of solution of the task, mostly copied material which includes a statement of its relation to the task, or are a broad general statement of the solution.

Score 2 A response which indicates that the subject has a basic understanding of the task and gives a relatively complete response to one part of the task. This would include a response which shows the relationship between two basic aspects of the task or which shows extension of one element well beyond the information contained in the passage. A response in this category may contain erroneous assumptions based on illogical interpretation of the reading material.

Score 3 A response which shows considerable extension of the material contained in the reading passage. The response must clearly indicate the relationship between two or more elements of the task, or in those items which require a digest of the material,

must cover all major aspects of the passage. Unlike lower level responses it cannot contain erroneous interpretaions of the reading passage. However, it may contain some irrelevant material.

Score 4 A response which shows considerable extension of the material in the reading passage and combines this extension with a novel application of prior knowledge. The response must clearly indicate the relationship between all parts of the task and unlike lower level responses cannot contain irrelevant material. In those items which require a digest of the reading passage the product must indicate that the student has combined several specific elements into general elements which when combined digest the entire passage.

Definition of Evaluation[1]

"Evaluation is defined as the making of judgments about the value, for some purpose, of ideas, works, solutions, methods, material, etc." (Bloom, et al, 1956, p. 185). These judgments involve the use of criteria which may be determined by the student or given to him. Thus evaluation differs from the other behaviors defined in the Taxonomy in that evaluative behaviors must be based on a definite criterion.

Most judgments made by an individual are quick decisions which are not preceded by a deliberate consideration of the different aspects of the thing being judged. These judgments are more properly labeled opinions since they are not based on an objective criterion, and as such cannot be properly labeled evaluation as defined in the Taxonomy. For the purpose of scoring items, only those responses which show definite evidence that the student based his judgment on a criterion, stated or unstated, should be classified as evaluation responses.

Evaluations may be of two distinct types in that they may be based on internal or external criteria. Evaluations made on the basis of internal standards. . . "are for the most part concerned with tests of the accuracy of the work as judged by consistency, logical accuracy and the absence of internal flaws." (Bloom: p. 186). Evaluations

---

[1]The material in this section is based on the Discussion of Evaluation in the Taxonomy of Educational Objectives Handbook I, Bloom, et al.

made on the basis of external criteria are made using standards arrived at by considering the ends to be served by the thing being judged. Evaluation, as measured by the Taxonomy tests require judgments on the basis of external criteria.

Response Characteristics

For the most part, the items will present "facts" concerning authors of the reading passage with the students task being to evaluate the effect such facts would have on the authors qualifications to write such an article. The standards on which the judgments are made must be supplied by the student in that he will be required to set up a hypothetical author in his own mind and relate the "facts" to this hypothetical author. Naturally the qualifications of this hypothetical author need not be stated by the student, however, evidence of such a construct will be evident in the way the "facts" presented are related to the qualifications. Scoring of the responses will therefore be accomplished by an evaluation of the extensions made of the"facts" presented. In other words the evidence of a criterion will be determined by whether or not the study makes an extension of the facts.

On one form, Stages of Economic Growth, the items pertain to the validity of the economic theory presented in the article. The process in this test is the same as the others except the criteria is the applicability of the theory in the instances suggested by the "facts" of the items.

Evaluation items are scored on a scale ranging from 0 to 2. The general characteristics of the responses in each category are presented below.

Score 0 Does not relate statement to criteria. Indicates that the student did not accept the "facts" presented in the item. A simple statement of opinion with no statement indicating basis of judgment. No indication that the student was able to formulate a relationship between "facts" and qualifications or formulate a basis for no relationship. Student missed an important implication of time or sequence. Student missed important "fact" contained in reading passage. (This is especially true of the items in stages of Economic Growth). Student generalizes accomplishments in one area to accomplishment in all areas.

Score 1 Shows basic relationship between "facts" given

in the question and relates them to criteria. Logically extends basic implication of "fact" and states opinion on the basis of this extension. States criterion used and states how the "fact" given relates to the criterion (this type of answer frequently found for "no effect" answers). Attempts to answer items without extension of the "facts" given and arrive at no effect due to insufficient information.

Score 2 Makes logical extension of "facts" presented and arrives at a judgment based on a resolution of conflicts between the extensions. Acknowledges information necessary to form an evaluation is missing, makes a supposition plus extension of "facts" presented. Translates entire article into a general category and translates "facts" presented. Translates entire article into a general category and translates "facts" given into specific attributes and makes an evaluation based on these translations (this type of response is common for items concerned with profession of author). Indicates judgment based on global aspects of the "facts" but points up assumptions basic to this judgment.

SECTION III

Scored Sample Responses

Atomic Structure

Item 81. A chemist claims that he was able to make a sample of xenon, a member of the Zero Group, combine with fluorine. Briefly explain what might have happened to produce this result.

Objective: The student should recognize that the chemist must have made an error and had an impure sample or that some powerful process was devised to force the reaction, because according to the article Xe will not combine with anything.

Examples Score 0

It will explode.

If the chemist had combined an inert element of the zero group with fluorine, there would have been no reaction.

He just discovered a way to make them unite, but it hardly seems possible after everything else that was tried failed.

They will not unite because the article says that even Fluorine, most violent of all the non-metals, cannot shake these hermit elements out of their inertness.

Examples Score 1

Probably extreme heat or pressure was exerted upon xenon, because zero groups do not normally combine.

Since the xenon was a gas it could have turned into a liquid and then combined with fluorine.

Evidently, the element he thought xenon was another element.

Perhaps he could produce a xenon ion that would combine with fluorine whereas the atom would not react with fluorine.

The only possible means of making an inert gas combine is to eliminate one of the electrons, which would make the gas unstable. This may happen in an atomic bomb or other nuclear force.

Examples Score 2

An electron of the xenon has been either gained or lost before the trial. The xenon, then is not xenon, as the chemist believes.

He probably had to use an extreme high temperature or a lot of pressure. He might have been able to make it by combining two elements, or by taking away or adding protons or electrons from a higher or lower group.

The sample Xe must have been impure and the fluorine combine with the impurity. This is ridiculous because the chemist would control his experiment better than that and never make sure a claim.

The Xe was probably a weaker member of the zero group and a small quantity was mixed with the fluorine which was strong and a greater quantity was used which enabled the two to combine.

Examples Score 3

Table 2 shows that there are missing elements at the end of the chart. He could have discovered one of these elements and used it as a catalyst, which would combine with both the fluorine and xenon and form a three-element compound.

First of all, assuming that the chemist was qualified and hadn't made a mistake in his work or allowed impurities to enter the reaction, he would have to try to combine them under very unusual conditions. This might be possible in an atomic reaction.

There must have been an impurity mixed up with the sample of Xe, something that would combine with fluorine. He probably just thought the xenon to combine he might have used an accelerator, or some atomic device with lots of force and power to make them go together.

A sample of Xe might have been placed in an atomic accelerator and bombarded by a mass of electrons forcing one electron upon the Xe. Thus, changing it to a difference valence like cesium maybe, making it able to combine with fluorine, because xenon as we know it won't combine with anything, so he had to change it somehow.

Examples Score 4

Since there are many new methods being developed in science, he may have found a way to make xenon gain or lose an electron. Possible he could have created a high positive magnetic field in an area so small that it would pull off only one electron from the xenon, thus causing xenon to be an ion and unstable.

He took Xe and F_2 gas in the ratio of volume of 1 to 5, and placed it in a nickel cylinder, heated to 400°C for about five hours, under very high pressure so the electrons will get loose, quickly-cooled it in a dry ice and alcohol bath, and crystals of XeF_6 formed.

This reaction could have occurred by either of 2 processes. A real tremendous catalyst could have been used or intense conditions of heat and pressure during the experiment. Since chemists have probably already attempted to combine all known elements with Xenon, a catalyst was probably no good so he most likely used extreme heat or pressure. Since much new equipment, like generators, accelorators, atom crackers are being experimented with, probably one of these new machines was used and produced a higher charge in the stress on the atom than was previously possible.

Item 82. The article states that Mendeleef started his table by arranging the elements into seven groups. How might his results have differed if he had started by arranging the elements into seventeen groups?

Objective: These are two approaches the students may take in answering this item. The negative view is appropriate when the student realizes that with 17 groups Mendeleef would not have had the periodicity and the vertical grouping. The positive objective is justified if the student realizes that the modern table has 17 groups and that had Mendeleef used this arrangement in the first place, it would have been more

complicated but he would have had a place for the inert gases and would have saved the work it took to derive the existing table. No matter which approach is used the responses should indicate that the number of groups is determined by the characteristics of elements, not by the desire of the chemist.

Examples Score 0

The groups would have more elements in them and the table would not be as complete.

There would have been fewer elements in each group and the way the grouping was set up would be different.

Such an arrangement would not have had any great difference except in the grouping of one element under a similar element.

He would have had too many groups which would have fouled up his theory on groupings and experiments with the elements.

Examples Score 1

He would have one, two, or three groups containing elements--similar to one another. The table would have been much more difficult to analyze and to recognize unknown elements.

His table would not have shown him that the properties of elements were periodic functions of their atomic weights. The table would not have been as orderly and revealing as it is is now.

When he arranged the elements in groups of 7 and brought sodium down under lithium, he could see the resemblance between the elements. But if he had brought the elements out to seventeen, Li and Na would be far apart and you wouldn't have groups according to physical similarities.

Mendeleef would have had more blank spaces in the 17 groups. The elements which had similar properties would not have been arranged correctly and the atomic weights wouldn't have been in order.

He would have come up with a table similar to the one we have today which has seventeen groups. He would have had to predict more elements and possibly the table would not have had such acceptance as it did due to the great missing parts.

Examples Score 2

His table would have just been an arrangement of chemicals by atomic weights since similar elements would have been placed in different groups. He might have attempted to arrange them by atomic number. If he had done this he probably would have discovered that the 0 group was completely missing.

He probably would have noticed that there were four groups in which three elements seemed strangely alike and that there were three groups in which only 2 elements seemed alike. He probably also would have noticed that this same order was repeated in each group of 17.

The elements could then have been separated into more finely graded groups. The similar elements could then be closely placed. This would make several groups of elements closely related but would allow for better placement of newly discovered elements.

Since the actual grouping of the elements is based on the physical and chemical properties of the elements in general, the number of initial groups would not matter. He would have combined like elements into groups, then having noticed the similarity of various groups, he would have combined them also.

Examples Score 3

If Mendeleef had arranged his elements in 17 groups, he would have had to look for more minute differences between the elements. That's why two elements almost exactly alike could be put in different groups. Also the atomic weights and activity could be placed wrong with so many groups and he wouldn't have seen the periodicity in the vertical groupings.

The elements that are now in "a" and "b" columns of the same group in Table 2 would be in different groups. This arrangement would also have left 3 vacant groups for the inert gases and similar metals. Thus, he would have a weaker chart when trying to find order in the existing elements but would have been a better chart when considering the fact he could have predicted 2 unknown groups.

Examples Score 4

According to the article Mendeleef grouped the elements according to their properties. He didn't start with a pre-conceived number of groups therefore he probably originally would have had groupings based on any significant differences in characteristics and would have realized his error and combined groups using more general characteristics.

If he had chosen 17 groups he would have seen into the problem of where to place an element that had characteristics of more than one group. He could have done this on the basis of atomic weights, but since atomic weights says little about the relationship between various elements he probably would have decided to use another characteristic such as the ratio in which the elements combine as a basis for his chart classification. Thus probably coming out with 7 groups.

Item 83. Mendeleef predicted elements to fill three of the gaps in his table, yet he failed to predict elements to fill the remaining gaps. Why do you think he did not predict elements to fill these gaps?

Objectives: The student should realize the bases by which Mendeleef was able to predict certain elements. The ones he predicted were in the middle of known elements and he could reliably predict them by the weights and characteristics of the surrounding elements. The ones he didn't predict were massed together in groups near the bottom and he had nothing to go by.

Examples Score 0

He couldn't find a suitable element to put into the gap.

He wasn't certain of the properties they would have.

The three he predicted are known to exist but their atomic weight wasn't known. The other elements hadn't been discovered yet.

He wanted to leave something for other scientists to discover.

Examples Score 1

Because he didn't have enough information about the properties and weights of the undiscovered elements.

The unfilled elements are in the high energy phase and are not surrounded by other elements of the same characteristics.

The three gaps he did predict were either preceeding or following common elements and it was logical that there would be similar common elements, while the other gaps were not near so common elements.

I feel he must have died before completing his works. This man was so great at predicting he only predicted when he felt he was very sure of what he predicted. He might have been in doubt.

He didn't have enough information to predict elements to fill the remaining gaps. If he had predicted any more elements he would have only been guessing and not setting forth a sound workable theory.

Examples Score 2

Since a lot of these gaps were surrounded by unknown elements, he couldn't accurately predict the characteristics of the missing elements. Also, these gaps occurred near the end of the chart and Mendeleef had no way of knowing how long it would e v entually be.

He did not know at that time that elements, beyond these that were known to exist, would be discovered. However, he could see that the gaps between known elements could be filled if techniques were developed to isolate them.

Because he could not be sure if there was a gap. He used atomic weights to get the position of elements and probably if there was a big enough gap between known elements he could be safe in assuming that there was another element here. But if there wasn't much of a gap he couldn't be sure whether it was just inaccurate data or really an element there.

Examples Score 3

The three he predicted were sandwiched between elements he knew well; thus he was well qualified to make a prediction

of what must come between. The other gaps are usually gaps of two, three, or four elements in a row, which certainly leaves much too wide a range of atomic weights, etc., for any degree of accuracy.

He did not see similarities which might lead to the new elements. These gaps also appeared under other gaps at times. Some elements which would be in group I, for instance, and be related to Cs, might not fit characteristics of period 4. Contradictions like these could be problems. There are the zero elements which he could not have seen for the above reasons, and perhaps others.

Mendeleef had an idea of what the three elements he predicted would be like because of the positions they held between already known elements. So, according to his table he knew what they would have to be like. But he had places where the gaps covered many elements so they didn't have a place between too well known elements so he couldn't predict them just by looking at his table. Research would have to be done to discover them.

Examples Score 4

Mendeleef failed to predict the remaining elements because of the technical limitations in scientific methods of his day. Even with all of the technology at the disposal of current chemists some elements have not been isolated. We should bear in mind that his work was far in advance of his day and the limitations of his work were due more to technology than to his ability. Another reason for his failure was that he used Atomic weight; for had he used atomic numbers he would have seen gaps by just counting.

Item 84. Write a brief history of the development of the periodic tables. The article will provide a basis for this history, but you should supply events which could have occurred in addition to those given.

Objective: The objective is for the student to make up a history of the periodic table. In doing this he must fictionalize rather than just to pick out and list the facts as they are stated in the reading passage.

Examples Score 0

Mendeleef began the idea of completing a chart to list the

elements in relation to one another. His chart has been filled in with new chemicals that were discovered after his death. Many forms were made and different layouts of chemicals are available.

Mendeleef discovered the periodic table and there have been lots of improvements since he died.

Mendeleef arranged the elements into seven groups according to their atomic weights. When he finished he noticed the elements were periodic functions of their atomic weights.

Merdeleef created it and it was immediately accepted and praised the world over as a masterpiece of genius. It has been improved on since.

Examples Score 1

The first recorded periodical table occurred in 1870 when Mendeleef discovered the natural order of the elements. At that time, however, several elements were not known and had to be left blank. Since that time most of them have been discovered and filled in. Fairly recently (1957) a new periodic table, based on atomic rather than weight has been devised.

Somebody realized the organization of the chemical elements by characteristic qualities. This led them to place these in some orderly form which led to the making of a primitive periodic chart. From this chart they realized other characteristics that were present and continued developing the chart to its present form.

1 - First discoveries of elements.
2 - Mendeleef supplies order to known elements by grouping them according to similarities.
3 - Further advancement and discoveries of new elements.
4 - Preparation of the modern periodic table.

Mendeleef was the first to develop a periodic table. He arranged the elements into seven groups but he left many gaps. Later two Englishmen came along and found the inert gases and called them the zero group. They could name the elements that Dmitri predicted. Current charts show atomic numbers, atomic weights and familiar characteristics.

Examples Score 2

Dmitri Mendeleef was the first chemist to list in some kind of order the elements. Up til his time scientists were just discovering the elements. Mendeleef, by placing the elements in order let us see the relationship between all the elements. Of course, during his time all the elements were not known, but he knew they would be found so he left room for any new ones. Later, Cavendish and Moisson found the properties of the inert gases which they put into proper order. Each year brings us closer to what we feel might be a complete periodic chart.

Dmitri Mendeleef found elements with similar properties and arranged them into groups. By noting consistencies combined with further study, he completed a table of known elements. He also predicted elements and left spaces for others, using known consistencies as a guide.

Further research by later scientists found the predicted elements and those which filled the empty spaces. A few changes were made in the table arranging elements to atomic order instead of weights. Energy levels and other known facts were used to improve the table. A chart of electron configuration is now used.

A man named Mendeleef became interested in the atoms of all the elements. He decided there must be some relationship between all these elements. He began by experimenting until he had placed them in a different order according to weight and he lined them up in groups of similarly looking and acting elements and found a remarkable pattern. Through the years he and later scientists found out a good deal about new properties and new elements, and Mendeleef's chart had to change again to fit all the new elements and the ones they made, and because they were using atomic numbers for a base now, and not atomic weights anymore.

Examples Score 3

Through history, as man began naming elements, he discovered that some resembled others in certain ways. Mendeleef recognized this in a more profound way, and organized the known elements into groups according to their similarities and atomic weights. Surprisingly enough of the relationship between the weights and the group structure held, even with the discovery of new elements and so the first periodic table evolved. The table has been continually used as a basis, with other facts

added on as research and discoveries developed-such as protons, neutrons, electrons, atomic weights, and orbitals.

Examples Score 3

1800 - Crude knowledge
1810 - Mendeleef born
1820 - Work progressed slowly in chemical fields
1840 - Introduction of steam engine, study of elements and fuels, new discoveries
1860 - Work on periodic table by Mendeleef
1900 - Mendeleef's work taken up
1920 - As technology advances, table is filled in
1945 - Work on the Atomic bomb helps fill in chart
1950 - Rocket age introduces new work on chemicals and elements
1960 - Synthetic production of new elements

Dmitri Mendeleef defined the weights and numbers of the atomic elements. Then placed each into one of seven groups. When the list was completed he saw that a definite pattern, with few contradictions emerged. He had discovered that the properties of these elements were periodic functions of their atomic weights . Also the uniting methods showed a definite relationship.

Mendeleef corrected the contradictions going against those who had given the weights and numbers to the atoms, this being heresy in the strictest sense. His table corrected, he attempted to fill in the blank spaces of his table-his predictions have proved correct. However, some blanks could not be filled in, these were later filled in with the zero group elements, discovered by two Englishmen, and later seventeen other elements were found, totaling 86.

Today there are many more elements which were discovered through the years, and the chart is now based on the weight of carbon instead of oxygen, these are the only changes.

Examples Score 4

Mendeleef began the periodic table by arranging the elements in rows of seven according to their atomic weight. Perhaps he knew nothing then of atomic numbers. With this table he predicted the discovery of elements which would fill the blank spots in his periodic table. Then the inert gases or zero group were discovered. With this

discovery it is possible that chemists could believe still other elements had to be found. When atomic numbers were found and related to the previous tables, they probably noticed that the atomic numbers were closely related but that's where the atomic weights showed no possible gaps between elements the atomic numbers did. Then chemists probably tried to find these elements suggested by the missing atomic numbers. Also, with discovery of a few new elements which fit into the atomic number gaps and not the atomic weight gaps, it became necessary to change the arrangement of elements to the order of atomic numbers.

In the beginning, people knew that there were things called elements, and suspected maybe that they combined with each other, but they just went about with their chemistry in a haphazard manner. Then, along comes this Russian guy who decides to make an orderly list of all these elements all over the place. So, he makes a list of elements according to their atomic weight and, What ho! it turns out to be even more orderly than he suspected. At first people say, "So what?" but then eventually they are convinced, and some people become fascinated with the idea that elements have order. After a while, they come up with this "periodic table" which becomes accepted all over the world.

One day these scientists were heating helium and somehow discovered that the little electrons jump into higher energy levels and what's more, they've got orbitals instead of shells. This calls for rewriting of the periodic table, so they throw out the old one and come up with the "electron configurations of the elements," based on new discoveries and this chart will do very well until someone discovers something else which will change it.

First classification of elements which compose the earth into four basic things; air, fire, water, and earth, primitive peoples are extending far into civilized scientific beliefs.
-Medieval times when the alchemists searched for the magic formula which would turn anything into gold.
-Discovery of these simple basic elements, gases of hydrogen and oxygen, advancement of methods and theory as pure science.
-More elements found and identified, tentative theories of classification
-Mendeleef comes along and revolutionizes classification of the elements with periodic grouping by atomic weight.
-Predictions of unfilled and more gaps filled in due to advanced techniques in research, whole zero group found
-Radium and radioactivity found by Curies
-Synthetic production of elements by atomic bombardment
-Further development in atomic research during and after World War II

-Age of rockets, emphasis on light weight compounds, study of atomic structure in infinite detail, study of fuels and metals for satellites and rockets, new characteristics and means of classification explored, and possibility of new elements in the outcome.

Item 85. Suppose that you were the first chemist to visit another planet. What steps would you take in constructing a periodic table of the elements found on that planet?

Objective: The student should recognize the steps taken in evolving the current periodic tables and specify how these steps would be used in organizing a new table. In his answer a student may also postulate similarities and differences in the elements of the earth and new plants and formulate plans for testing the assumptions.

Examples Score 0

You would have to gather elements in the three different forms (solids, liquid and gas) and analyze them.

I would find out everything I could about their atomic structure and then follow Mendeleef's cue.

If you suppose I am the first chemist to visit a new planet, I suppose that by then I would know how to start a new periodic chart of a bunch of unknown facts.

Test theirs against ours; eliminate and add those that differ.

Examples Score 1

Find physical and chemical similarities, find atomic weights. Calculate the periodical repetitions of given characteristics. Then set up each chart according to that interval. It could even be 22 instead of 8.

Find the weights and combing properties of each element found, and arrange them according to similar combining properties and progressive weights.

I would try to fit the elements there to the periodic table

we have here, for there is no reason to expect they would be radically different.

Spectroscopic analysis of the materials found on the planet would be the earliest method of identifying them.

Examples Score 2

I would take samples of everything I could find into my lab. Then I'd perform electrolysis on my samples and turn all compounds into their respective elements. Then, I'd take 1 cc of each element I had and weigh it. From these weights I'd determine which would be the heaviest, lightest, etc. I'd assign the lightest one a number of 1 and use that as a basis of assigning weights.

Experimentation is the first step. I would try to find the basic elements in the air and surroundings. I would try to compare their properties with the known elements. After finding out what I thought to be fairly accurate amount of properties of many elements, I would place them on the chart according to their atomic number. I would leave vacant any area which I was not sure of for further findings.

First, I would set up a base element, one that would be used to determine the weight and number for the rest. All with similar characteristics would be placed in one group. Such things as if they tend to become negative or positive ions would be observed. After all information possible was received, then the similar elements would be placed in a group running vertical. Energy increase potential would be on the horizontal line; further to the right, the higher the energy.

List the elements in some order of sequences according to similar properties. A comparison to known earth metals might be useful. Readiness to combine, weight, physical properties would all be characteristics which would influence their placement on a table. There elements appeared unrelated, gaps could be left in hope that later elements could be added.

Examples Score 3

-Make a list of known elements
-Test them for similarities in their behavior
-Begin placing them in groups according to the number of general and recurring characteristics

-Discover the pattern of the electron, proton and neutron distribution
-Once in groups, they must be arranged in order of increasing energy
-Space for as yet undiscovered elements must not be omitted because some elements will be difficult to find.

If the elements were vastly different from these found on earth, I would first try to determine the new elements' atomic structures. This is done by attempting to combine them with elements of known structure. I would record all the results and compile the characteristics of the elements. Also, the atomic weight could be determined through a comparison with carbon.

After all facts have been determined about the elements, I would attempt to find the order of the chart by placing the elements in order (by atomic number) and putting them in columns by characteristics.

I would first gather all the elements that I could. I would then choose one and react all others with it grouping them according to the similarities of reaction. I would then proceed through checking properties, weights, etc., to group them. Extensive tests with other elements would be necessary to assure any accuracy because perhaps most of them wouldn't even react with the first one I chose. I would have to decide on a basis for my beginning separation and it would be the reactions. Later I would go on to check physical properties, etc., and make finer distinctions.

Examples Score 4

The first thing I'd do if I were the first chemist on an uninhabited planet is to check around and get some samples of all the elements I can find. I'd identify the ones I knew that we have on earth and then isolate the ones I didn't know and try to classify them on standards of the similar earth elements. After determining general characteristics and making groups of the various different types of elements, I would make tables of their atomic weight and numbers and decide which would be most effective as a basis of order. I would have to leave a few spaces for the ones I couldn't find, and if I have really found the pattern of occurring, then my table would fill out just like Mendeleef's eventually did.

-Get samples of all available elements

-By experiments find out:
  combining ability
  conductors of heat and electricity (metals, etc.)
  normal physical states
  atomic structure
  energy levels
-Then finding atomic weight and going by discovered properties, separate them into groups
-Determine periodicity
-Using atomic number which is earth basis put them in order by periods horizontally and the groups should line up vertically if it is right

First of all I would have to start my plans before leaving the earth. I would get all of the available astronomical data for the planet in order to determine what types of elements I could expect. Also since many scientists feel that all matter has a common source I would devise means for detecting elements for which there are gaps in our table and also for those higher than our table. After arriving on the planet I would begin by systematically identifying the elements found on earth, most or all should be identified. Then, I would try to identify elements 104, 105, 106, etc.

Evaluation Sample Responses

Item 86. He is a professor of Chemistry at an Ivy League University.

Examples Score 0

Qualified. He should know his chemistry.

Not qualified. Just because he's a professor of chemistry doesn't mean he's studied the periodic table.

No effect. He could have studied the chart anywhere.

Examples Score 1

Qualified. To become a professor he would have a background in the periodic table since it is necessary for the study of chemistry.

Not qualified.

No effect. He probably has studied the periodic table but

he may not have studied the history of chemistry enough to write a out it. He might have specialized in oil and its properties.

Examples Score 2

Qualified. This man is not only a professor at a university but at a highly respected (Ivy League) university. This means that he must have a good background in research and therefore he would know where to find the information to write on the periodic table.

Not qualified.

No effect. As a professor at a good university he probably has the background in chemistry and the research experience to find information to write an article, but can he write? There is no guarantee he can get across what he wants to.

Item 87. He is not a member of the American Chemical Society.

Examples Score 0

Qualified. He is a philosopher.

Not qualified. He is not a good chemist then.

No effect. A lot of authors are not members of the A.C.S.

Examples Score 1

Qualified. Because to be a member of A.C.S. you have to be a chemical researcher and this person was an historian.

Not qualified. Because the American Chemical Society is an important organization and all leading American chemists would belong to it.

No effect. Since most of this article is about a Russian maybe the author is Russian too and is a member of the Russian C.S.

Examples Score 2

Qualified.

Not qualified. The early parts of the article deal with history which anyone could read from a book, but the last part deals with modern things. You would have to know a lot of chemistry to get all the facts right in a short article and since the new chart is by an American I assume the author is American and if he is a qualified chemist he would be a member of A.C.S.

No effect. The persons qualifications should be judged by the accuracy of his writing not by his club memberships. Maybe he is a smart college student but doesn't have the degrees to get into the A.C.S., besides for all I know the A.C.S. may be a group of crackpots.

Item 88. He discovered the element Am.

Examples Score 0

Qualified. This shows he knows what he is doing.

Not qualified. He would not know history.

No effect. Maybe it was an accident.

Examples Score 1

Qualified. In order to know he had found a new chemical he would have to know a lot about how the elements fit together and how the chart was made.

Not qualified. Because the author would have to know about history of chemistry and if he discovered Am this means he is a researcher and wouldn't be interested in history.

No effect. A lot of discoveries have been made by accident by people who don't know what they have found, but have been told what it was by someone else.

Examples Score 2

Qualified. He must have studied a lot of chemistry in order to find a new element. Since Am was not on Mendeleef's chart, he probably had started by a careful study of the chart and found out what characteristics the element must have. He would have had to study the later charts too, so he would have to know where it fit, so after all this study he would have the information.

Not qualified.

No effect. Because this doesn't tell if he knows about the history and all of the changes. He could have found this information by looking in some books. Discovering an element doesn't say he looked in history books. Besides he must be able to write in simple English for high school kids and a lot of scientists can only write with a lot of X's and ='s.

Item 89. He recently wrote a high school chemistry book.

Examples Score 0

Qualified. A high school chemistry book contains the periodic tables.

Not qualified. A high school chemistry book isn't written on a high enough level.

No effect. A high school chemistry book may or may not contain a chapter about the periodic table.

Examples Score 1

Qualified. The very fact that high school chemistry books contained section on the periodic table means he is able to write about them.

No effect. High school chemistry books have more than one author; maybe this author didn't write the section on periodic tables.

Examples Score 2

Qualified. Our chemistry book was written by a Ph.D. at Harvard and contains a whole chapter on the periodic table. I guess for a book to be accepted the author must be qualified. If all high school chemistry books have a section on periodic tables the author must be qualified to write on them.

Not qualified.

No effect. He might have written the book but maybe it was so bad that he couldn't get it published.

And even if it were published that doesn't say the author is qualified.

Item 90. He constructed a modern Periodic Table.

Examples Score 0

Qualified. Then he must know about the old ones.

Not qualified. Because he didn't believe the old ones were any good.

No effect. Maybe it was a table of wars.

Examples Score 1

Qualified. Because he would have studied the old ones to see where they were right or wrong.

Not qualified. Because he would think his was a lot better and would find fault with all the old ones to make his look good.

No effect. Because it would have no effect if he was using a different way of grouping the elements.

Examples Score 2

Qualified. In order to construct a modern chart he would have to know a lot about chemistry and how elements are alike and different. His history would be one which would show what different people used to classify elements and how their work provided for the clues to undiscovered elements.

Not qualified.

No effect. Maybe he is trying to make a table people without much chemistry can understand--like me, for example. In this case he would take a complicated table, like table 2, and make it more simple, but wouldn't have to study all of the old ones.

Item 91. He is a teacher of high school chemistry and does not use a Periodic Table in his classes.

Examples Score 0

Qualified. The periodic table is too complicated for high school students; he is wise.

Not qualified. If he doesn't use the table, how is he qualified to write about it?

No effect. Just because he doesn't use it doesn't mean he doesn't know about it.

Examples Score 1

Qualified. A teacher of high school chemistry must have a college degree and therefore would know where to find the information to write about the periodic table--if he didn't already know about it.

Not qualified. Since the periodic table is the basis of chemistry, his not using it indicates that either he doesn't understand it or he doesn't realize its importance.

No effect. Maybe not using the periodic table is just his method of teaching chemistry; maybe he thinks it's too hard for high school students.

Examples Score 2

Qualified. First of all he must have a degree in Chemistry to be teaching it. The fact that he doesn't use the table in his class doesn't mean he doesn't know about it; maybe it's just his teaching method. With his background in Chemistry he could find out all about the subject and write about it.

Not qualified. To write about something you would have to believe in it and therefore you would use it whenever you could. This teacher obviously doesn't really believe in something which is very close to the basis of all chemistry, so how could he write about it.

No effect. This statement could be interpreted two ways. There are good and bad high school chemistry teachers. If he is a good teacher, we can assume that he has a reason for not using the table; if he is a poor teacher maybe he doesn't know it enough himself to teach it. Even if

he were a good teacher, he still might not be qualified to write.

<u>Item 92.</u> He has a college degree in Russian History.

<u>Examples Score 0</u>

Qualified. Mendeleef was a Russian.

Not qualified. History is not related to science.

No effect. It doesn't say if he did or didn't study Mendeleef.

<u>Examples Score 1</u>

Qualified. Since the first table was by a Russian scientist, he would have known about him in history.

Not qualified. The study of Russian history wouldn't say anything about modern tables constructed in other countries.

No effect. This only says he probably studied a lot and read some about Mendeleef. If he was interested he could have written the article.

<u>Examples Score 2</u>

Qualified. In order to get a college degree you have to know your subject and be able towrite well. Mendeleef was a famous Russian, so he would have read about him. He could easily read about modern tables and write a good history.

Not qualified. Because to write a good summary on something you have to know a lot of facts. In trying to figure out the modern tables and why they were changed, he would get confused and write things that may not be true. In other words he could copy from hisjtory books, but could not figure out present day chemistry without a lot of study and a history major doesn't study much chemistry.

No effect. The fact that he studied Russian History doesn't say if he studied Mendeleef's chart or any others. The article says Mendeleef was a philosopher and history is more philosophy than science. So he would have heard of

Mendeleef, but may not have been interested in tracing his work to today or he may have if he was interested.

Item 93. He wrote the article for an assignment in a college English course.

Examples Score 0

Qualified. All you have to do is look up the facts.

Not qualified. A student isn't qualified because he doesn't have the background.

No effect. It doesn't matter if the article was written as an assignment.

Examples Score 1

Qualified. A college student has the ability to write well and knows how to look up information.

Not qualified. Most college students use encyclopedias for reference and this wouldn't be an adequate background to write on the periodic table.

No effect. Maybe he was a graduate chemistry student taking English for kicks.

Examples Score 2

Qualified. To get into college you have to be smart and know many things such as how to write so people can understand you. Also, you have to know how to use the library. A student would be qualified to write such a poor article as we read.

Not qualified. A student writing for an English class would be more interested in correct spelling and grammar than in what he is writing about. He might have even copied all of the facts from the encyclopedia.

No effect. There is no reason why a student could not be qualified to write the article. On the other hand, he could have copied the article from a book. He could be a Chemistry student specializing in the table.

Item 94. He was educated in England.

Examples Score 0

Qualified. The English schools are the best.

Not qualified. English schools are not good in science.

No effect. It doesn't make any difference where he was educated.

Examples Score 1

Qualified. Because the article says that two Englishmen helped discover the zero group. He would have studied about them and learned about the old charts.

Not qualified.

No effect. Where one is educated doesn't make any difference; it is what is learned that counts.

Examples Score 2

Qualified.

Not qualified.

No effect. England has a very advanced technology and has a lot of good scientists, so he has a good science education he is qualified, but if he studied something else he wouldn't be qualified so the fact he was educated in England has no effect by itself.

Item 95. He believes that only about half of the elements have been discovered.

Examples Score 0

Qualified. He may be right.

Not qualified. He is some sort of a nut.

No effect. He can believe what he wants to.

Examples Score 1

Qualified. This shows that he knows a lot about the elements and has searched all the old charts and found gaps nobody else has.

Not qualified. Because many people think other planets have a lot of new elements, but these men are not usually scientists.

No effect. Because this doesn't say anything about his knowledge of the history of Chemistry.

Examples Score 2

Qualified. The fact that most of the gaps are filled in the chart in the reading passage means that he has discovered spaces for a lot of new elements to go on the end. In order to do this he would have to know a lot about the old charts and know what information was used to make them.

Not qualified. A real good scientist would do like Mendeleef did--not predict more elements then there was good evidence for. This shows that this man doesn't understand much chemistry because his "new" elements are probably ions or compounds.

No effect. Because the study of new elements has nothing to do with old charts. I think chemists today work only with the new charts, so he might not know much about the old ones unless collecting them was his hobby or something.

Glaciers

Item 81. Suppose that a glacier is advancing toward a town and that you have been asked to estimate how long it will take for it to reach the town. Briefly describe the steps you would take to make this estimate.

Objective: The subject must recognize that this is a rate-time problem which will require physical measures plus consideration of other factors which would logically be involved. Such factors as seasonal changes, mass, terrain configuration, and their effects must be applied to the prediction.

Examples Score 0

I would estimate the speed, then be able to determine how much time there was til it reached the town.

I would keep accurate check on the amount of movement of the glacier, temperature and air moisture.

Find out how fast it's moving and divide.

Estimate the rate of accumulation and melting and compare.

Examples Score 1

First, figure the rate of speed at which the glacier is advancing. Second, find the distance from the glacier to the town. Third, find the snowfall, melting temperatures and conditions that affect the glacier.

Determine the speed of the glacier, taking into consideration the accumulation and loss the effect of temperature and calculate the time to reach the town.

I would place a line of poles across the glacier, some in the middle, some on the sides, and two on the outside boundaries of the glacier. In this way I would be able to chart the flow of the different parts of the glacier each day, and total movement in relation to the poles placed out of the bounds of the glacier.

Measurement of one day's movement of the glacier should be made. By dividing that measurement into the distance to the town, you will get the length of time it will take the glacier to arrive there.

First, I would find the average temperature, then a record of approximately how far the glacier has moved in past years, then I would go out and study the path it would follow to the town.

Examples Score 2

Since the glacier moves under its own weight, you would have to estimate its weight. The glacier is moving from a higher to a lower elevation so you must check the degree that the land slopes toward the town. Temperatures must also be considered, since a glacier will move faster during the summer months.

First, I would get a fix on the glaciers present position and then I would take a series of measurements and get the averages in advancement. During the time after and before, I would check the glacier at various points to look for cracks as other things that might cause slides or danger. Also, I would watch the weather and the amount of avalanche activity.

I would first see where it came from and how long it took to get here, and then I would see what kind of country it had covered and then see what kind of country it is going to cover. Make an average of its speed and you could roughly see how long it would take.

I would first find out what the annual summer temperature is for that area, then find out what the average snowfall is every year. I would then measure the run off to see how much is taken away. Then, I could estimate its speed and the time to reach the town.

Examples Score 3

If a glacier was advancing toward a town, I would:
1. Determine the rate and amount of snowfall in the area of the glacier and approximate the run-off.
2. Determine the amount of time it takes the glacier to go a certain distance.

3. Take the amount of miles to the town and divide.
4. Take into account the terrain, slopes, etc., as obstacles which would slow down or speed up its advance.

I would find out the average temperature of the town, average snowfall, the angle of the slope of the land, and the highest and lowest temperatures and the amount of rainfall. Then, I would calculate the amount of snow that would accumulate on the glacier, the run-off and the rate of the glacier. From this I would determine the final rate of the glacier and the distance from glacier to town, and finally calculate the time it would take to get there.

The first thing would be to get data on how much snow has fallen each year for the last several years and this year's prediction. Then, by finding out how much headway the glacier makes in winter and how far it retreats in summer, you can find its forward movement for several years. If you know the distance to the town, and consider things like sudden changes in weather and the different type of land, the estimate should be fairly dependable.

Examples Score 4

1. Obtain weather information from the weather bureau on conditions in the area and future conditions.
2. Determine the distance the glacier has already covered and then figure the average movement for a week, and mark the distance it moves for a given time period, according to the weather.
3. Make a study of the topography and probable path of the glacier, and factors which might alter its course.
4. Approximate what effect change of season could have if it were to take a long time to get there, a late winter freeze, or an early spring thaw, and check on past records of the glacier, because a sudden thaw could increase the rate by a huge amount and disprove your estimate completely.

First, I would consider the general climate. It would have to be pretty cold and have a lot of snow, or a glacier would never be able to come that close. But it might have hot summers and it would melt a lot. The surface of the ground would have an effect, too. If it were mountainous the glacier would have trouble going up mountains and would go down pretty fast. If it were a slow incline, up or down,

it would have less effect on the movement of the glacier according to the seasons. Such things as forests or cliffs m ight impede its travel, too. If it were little and going up hill it wouldn't go nearly as fast as a big one going down hill. So, after I had studied all this, I would calculate its present rate of movement, predict a reasonable future rate, and divide into the measured distance to the town.

Item 82. A scientist believes that advancing glaciers slide on a thin layer of water which lies between the bottom of the glacier and the earth. Outline a plan to test whether his belief is true or false.

Objectives: The objective of this item is to test the students ability to comprehend the difficulties involved in the most obvious solution of direct observation and to suggest a workable alternative. These alternatives may be in the form of special procedures or miniaturization of the problem.

Examples Score 0

A blockade could be set up to block the ice flow and any water build-up could be observed.

I. Find two glaciers:
   a. One glacier stable
   b. One glacier moving
II. Examine both glaciers:
   a. The stable glacier should not have a thin layer of water underneath.
   b. If the moving glacier does have a thin layer of water, make other comparisons of moving and stable glaciers.

Well, as the glacier moves forward it also decreases, but at a slower rate. This would always leave some runout in front of it, or a thin layer of water to move over, but doesn't mean it moves over it.

False, because the water would run off from the glacier. Although if there was a constant drainage of melted water to the bottom of the glacier it might be plausible.

Examples Score 1

Drill a core into the glacier and examine the portion where the ice meets the rock.

Drill a hole through the ice and study the core samples. Or, you could determine the temperature under the glacier by going in crevices, etc. If the temperature was below say 30 degrees it's unlikely there's water at the bottom.

Determine the path of the glacier. On the path, dig a station reinforced with a tiny hole in the surface. When the glacier passes over, you will see if water leaks in. If it does, that assumption is true.

If the glacier is not too thick, then pumps and drills could penetrate the glacier until the bottom was found. Then, put the pumps into action and if water is obtained from the bottom, the statement is true. I definitely believe this statement because that's the way a block of ice slides on a warm day.

Examples Score 2

1. Place an ice cube on a slight hill.
2. Make a terrain in miniature and establish similar conditions.
3. Watch action of ice cube.

Make your own model of a glacier, having the same conditions as a real one. Keep testing and observing. Do this many times.

A glacier is the same as a giant ice cube. You could use an ice cube to represent the glacier. You could see whether the statement was true for the ice cube, and if it was true then it could probably be true for the glacier.

I would somehow build two man-made glaciers and then put one with a thin layer of water over the earth and put the other one on plain ground.

Design a chamber that will have the same conditions as the glacier: air moisture, temperature, and snow fall. Devise a means of keeping them at the designated conditions. When a glacier begins to form, check for your conclusions.

a. Look at the glacier bed, both behind and in front of the glacier.
b. Construct a scale model and experiment, using the same times the glacier takes to move from one point to another, and use different layers between the earth and glacier, one of them being water.

Under controlled freezing conditions:

Set up rocky-sandy run
Start big ice hunk down
Note moisture and surface changes
I know ice will slide, but I also know that the bottom of ice surface might melt if it came in contact with a warmer surface than itself.

I would prepare a test in the lab and have the entire glacier visible at all times. This test would give me the necessary information as to whether glaciers slide on a thin layer of water. My test would be very real, because I would let a glacier advance under the same conditions it advances at present.

Lab Test:

1. Construct artificial "terrain" with glacier.
2. Lower temperature substantially.
3. Allow glacier to move down terrain and measure rate of descent.
4. Reconstruct terrain and glacier exactly.
5. Raise temperature to normal and time descent.
6. If higher temperature allows glacier to move readily, then the increased amount of water under it allowed it to move faster.

I would take a large cube of ice below freezing temperature with a lot of weight stacked on top of it and push it back and forth on a rubber carpet. There would be a powder water die on the carpet. If there was a liquid, the powder water would smear, if not it would not smear.

Item 83. The land over which a glacier has passed suffers considerable damage. Briefly outline why glaciers cause damage and describe the kind of damage they might cause.

Objective: The objective of this item is to present a situation in which the subjects may build upon implications in the article. The complete answer goes beyond the obvious solution presented by cursory examination, that of damage due to mass. It includes such things as topological changes, climatic changes, destruction of plant and animal life and the like.

Examples Score 0

Glaciers cause damage because of their tremendous weight, they smash everything.

They freeze everything.

They kill the people and animals.

Everything in the path of the glacier is torn down or smashed because of the weight of the glacier. The weight just crushes everything, all the trees and anything in its path.

Examples Score 1

The weight of the glacier can change the contours of the land a great deal, by flattening hills and forests, filling in crevices and forming lakes. Also, a glacier carries material along with it which it deposits along the way.

These glaciers are ice and would freeze the grass and vegetation on the earth where they slide. Also, they would ruin the soil so that no other vegetation would grow there.

Examples Score 2

Ice is hard and it contains small or large particles of harder material or rocks. It will "sand paper" ground and wear it. Loose material will be pushed along or packed, because of ice hardness, density, etc. Damage might be grooves in rocks, valleys and trenches where loose or soft material is removed. Trees removed, and miscellaneous material left behind when glacier recedes.

The glaciers pick up rocks and debris while moving the debris under tons of pressure, crushes scars, etc., everything it passes over. It's like a huge bulldozer going down a valley leveling everything in its way, and pushing the junk into more obstruction.

Glaciers are large, heavy moving pieces of ice. For this reason, they can cut trenches in ground, ruin crops, carry away soil, and even move large boulders. When the glaciers melt they can cause the formation of lakes, rivers, and ponds.

I. Cause
   A. Large moving mass
   B. Extreme cold
   C. Water damage due to run-off
II. Effects
   A. General leveling of objects
   B. Death of plant life
   C. Death of animal life

Examples Score 3

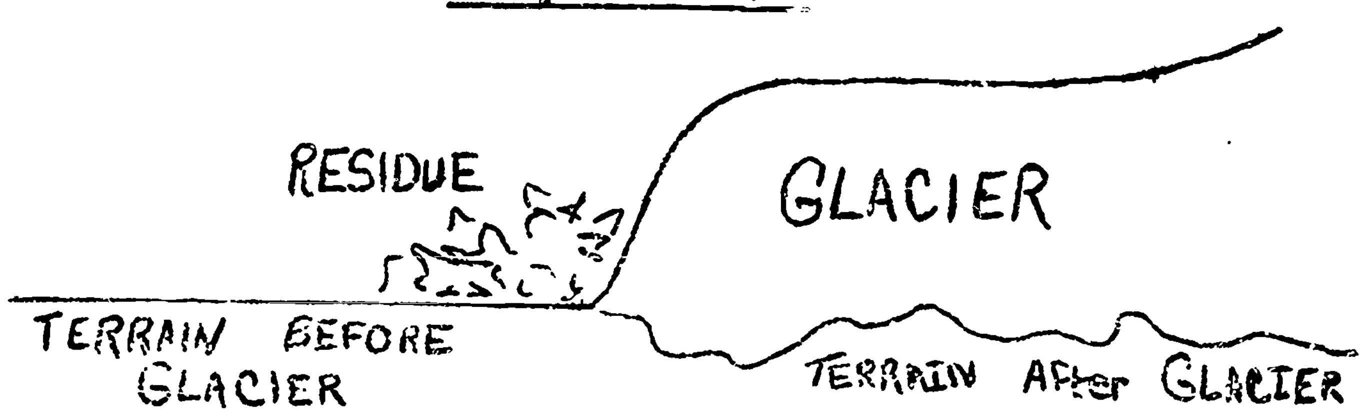
Glaciers do not go around objects, they go over them. Say, for example, a glacier overtook a small forest. The glacier would freeze the trees, then as it moves, carry them with it. The trees, rocks, etc., are dragged along the ground, tearing large holes in the surfaces.

Life in a glacier does not exist. After hundreds of years the ground is useless. River beds are often the result of glacier growth and shrinkage. Even after the glacier does receed, it will take hundreds of years to recondition the entire territory.

Glaciers weigh a considerable amount. Moving so slowly over land, it could cause ruin to plants and top soil. The run-off could flood a valley and kill animals. A glacier might go to a city or town and destroy millions of dollars worth of property and homes. It would take the land years to grow and mature enough to be able to cultivate or for grazing land, or just for plain forest area.

I. Glacier damage: Depending on what type of ground it moved over.
   A. Whether there were house and structures.
   B. If the terrain were smooth, gashes like the Grand Canyon would appear and mountain ranges would build up where the glaciers stopped. (Example)

Examples Score 4



Glaciers eliminate all life, both during and after passage. The climate conditions associated with glacier formation make it virtually impossible for plant and animal life to survive prior to glacier passage. The bulk and tremendous weight of the glacier removes the top-soil and humus, making it impossible for the ground to sustain vegetation.

The terrain itself is changed. Fusion is a constant problem

and much of the ground after ›sage is unmanageable and untillable due to debris. Af..r glacier passage, the weather conditions are usually still bad and floods as well as long periods of cold weather keep the territory from rebuilding rapidly.

When a glacial age occurs with glacier passage imminent, the climatic conditions prevent normal vegetation growth with an accompanying decrease in animal life due to starvation and migration. Erosion is a problem, both before and after a glacier due to this destruction of vegetation. The glacier itself exerts such a large force as it moves that valuable topsoils are scraped away, leaving nothing but unfertile subsoils. It takes nature hundreds of years to replace what is lost and sometimes this is never accomplished.

They freeze plants and roots. When they move, they dig up entire land surfaces. The water content is over-abundant and exceeds the amount the land can take in. Human beings and animals are unable to survive in the cold. When they melt, the runoff floods valleys and rivers, destroying more land and homes. The land may be frozen for many years, unable to reproduce growth.

Item 84. Briefly describe a method by which you could determine if the vertical thickness of a glacier is increasing or decreasing. In your method, indicate measurements you would take and how would you use them: Remember that glaciers move!

Objective: The student must recognize that the measurement of the vertical growth of glaciers involves more than drilling holes and measuring thickness. The complete answer involves not only mass movement but also requires a valid method of establishing a base point for measurement and keeping track of the base point, compensating for terrain configuration, the establishment of a logical plan for continuous measurement checks and the use of particular devices such as siesmographs.

Examples Score 0

Drill a hole every year

Look for cracks and measure their depths

See how far it moves

Examples Score 1

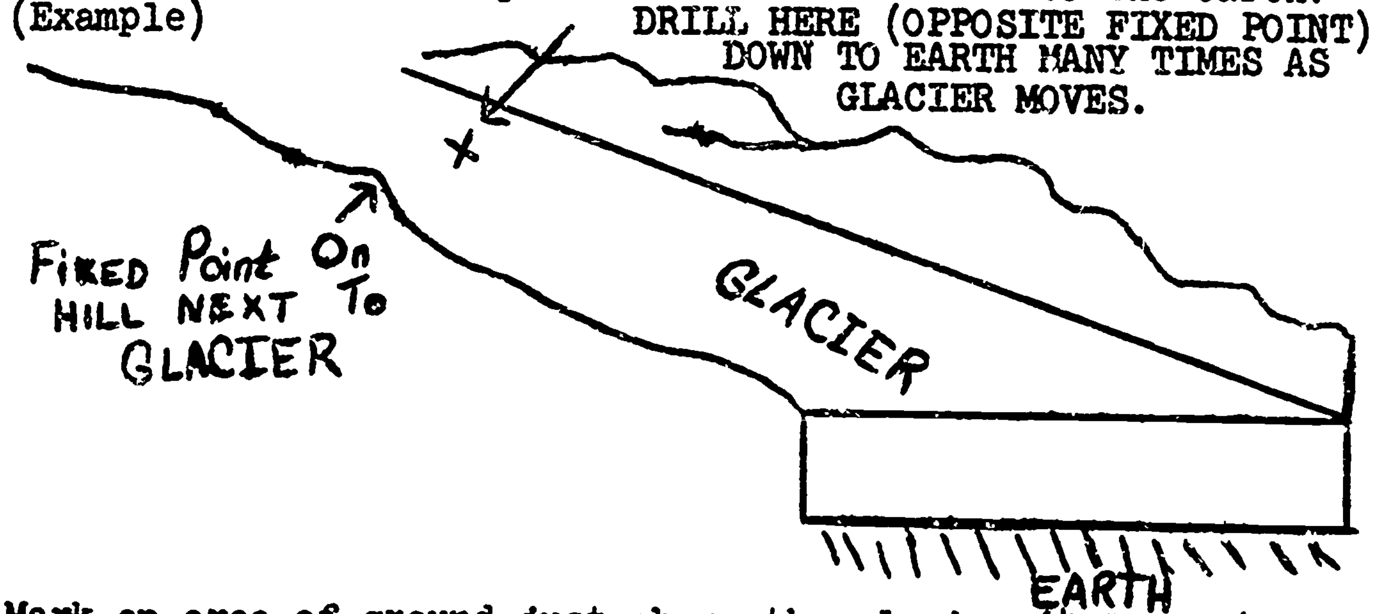
Take measurements of the thickness of a small amount of the glacier at different intervals and compare them after a certain period of time.

I would take a reading on its altitude by using an altimeter. I would take several of these at various times. I would then calculate the average height over a given time.

Just stick a pole down into the glacier. Mark the pole in feet and inches and as the thickness increases, just look on your pole to see how high it has risen.

Examples Score 2

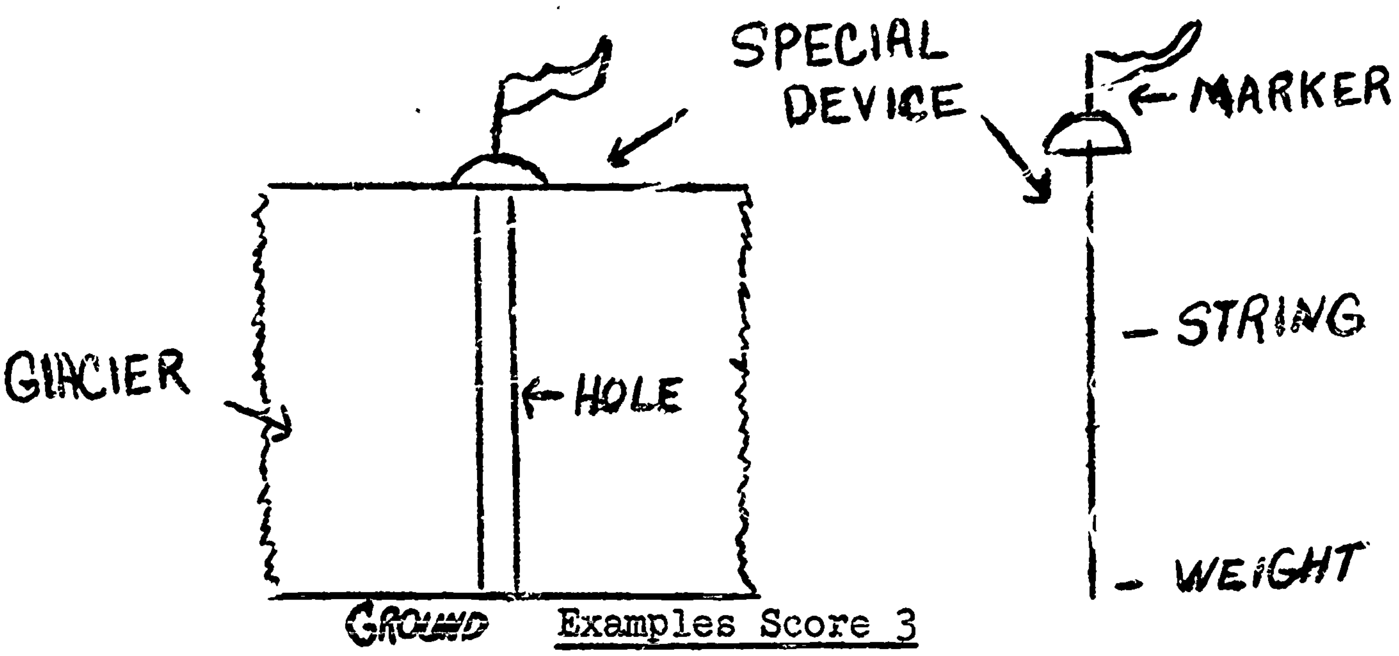
Pick a spot where you knew the point where the earth was. Line up a spot so you could stay on that point no matter where the surface goes. Just keep measuring and recording. You would measure the points from the surface to the earth. (Example)



Mark an area of ground just above the glacier (4-5 feet) and measure the distance the top of the glacier is from the mark. At pre-determined intervals come back and take another measurement.

1. Buy a 4,000 foot drill
2. Drill hole in glacier
3. Put special marker in hole
4. Check readings

(Continued on Next Page)



Examples Score 3

I would bore through the ice with an apparatus similar to those used to take samples of the ocean bottom. I would then insert a pole in the ice down to the level of the ground. This would have to be an extremely strong pole capable of withstanding great pressures and the shifting of ice. I would have the pole marked as to feet and inches and would also do this in several other places to make sure. The pole would move along with the glacier and measure any build-up.

I would calculate the height of the glacier by using trigonometry. Calculate rate of incline and the angle of inclination from the base of the glacier (using an altimeter to measure the air pressure) to the level of sea level and knowing the distance above sea level at the base of the glacier, could figure the height of the highest point, subtracting the height above sea level. I could do it periodically and the information collected the first time I could figure out if the glacier was increasing or decreasing, always considering the varying distances from sea level.

By placing a laboratory on the glacier's surface that would move with it, it would be possible to work better. From there the scientist could drill until he reaches bottom and take a section through the glacier to be brought up and measured. This could be done at intervals.

Examples Score 4

Assuming that all the necessary equipment was available, first mark the spots for measurement. Set in place marker

poles of some type. Use a drilling arrangment similar to that used by oil field workers. Set up observation stations to keep track of the marker poles as they move with the glacier. Take readings in the same general locations as the first readings to determine whether intervening snow falls have actually packed down enough to increase glacier thickness.

Use a seismograph to plot out the terrain features under the glacier. Mark spots for drilling or echo soundings. Repeat these measurements over a period of years to see if conditions have permitted the ice to increase or decrease. The places for measurement should be carefully marked on a map so that measuring parties can return to the same spots every year.

Select an observation point and relate it to some immoveable object. Take corings of the glacier to measure its thickness at that point. Wait a pre-determined length of time and take another coring and measurement. This new point will not be the same as the first point on the glacier since the glacier has moved, but will be the same point as related to the fixed object. Doing this at regular intervals could calculate the fluctuations in thickness.

Item 85. Suppose you knew that an advancing glacier would reach a town in 12 months. Outline a plan which could be followed to prevent the glacier from reaching the town.

Objectives: The objectives of this item are two-fold. First, the subject must devise a logical method for stopping the glacier and secondly, that most procedures would result in much ice of water to contend with and a method must be developed to dispose of it. Ample consideration must be given to the enormous mass of a glacier.

Examples Score 0

Move the town

Sell ice cubes

Haul it away on a truck

Examples Score 1

By increasing the melting rate of the snow (by salting the ice or treating it chemically) and then diverting the water flow away from the town, the glacier would melt.

First, I would try to detour it. Then, (if it didn't work) I'd use atomic weapons to move it, or to make it detour.

One of two things could be done in this case. You could either move the town or dynamite the ice. By dynamiting the ice you break it up into smaller pieces which will melt easier, come the spring thaw.

Examples Score 2

I would try to stop the glacier by using explosives. I would try to halt the progress by destroying a part of the glacier every few days.

I would try to stop it by using a system of heat or hot water to melt the glacier and slow its progress.

I would try to build a large "man-made" mountain in the path of the glacier in order to divert it.

Since rough ground breaks up glaciers, I would make the ground as rough as possible to break the glacier up. I would build some type of dam, and when the glacier got pretty near, I would use artificial heat to melt it, and let the water go into the dam.

Some way would have to be found to destroy the glacier. This could be done my melting it or using explosives. A chemical could be put on it, causing it to melt. Of course, the problem of the melted ice would arise and the water would have to be disposed of. By using an explosive, the glacier could be broken down and the movement stopped.

Examples Score 3

Dig a drainage ditch near the glacier to carry off excess water. You might run it through a hydro-electric plant, use it for irrigating fields, or for drinking water; then,

manufacture a big solar mirror and focus on the ice to melt it.

First, I'd find out what was causing the glacier to move, then if it were going to cause more movement, I would destroy the source. If destroying the source didn't work, I would plant explosives at several points in the ice pack. Then, let them explode and smash the ice. Then I would set up a system for removing the ice. If this didn't work, I'd try to push the ice into an ocean current, or, if there were no ocean, I would try to melt it with a heat source such as a la[illegible] magnifying type cover that would cause the sun's rays to [illegible] it.

In the year allotted you could erect a large coil a little in advance of the glacier. Behind this a ditch could be dug. Drop the ditch off to form a water-fall and put up a hydro-electric plant. Using an outside source, start electricity through the coil. This will require many millions of volts. Once the glacier begins to melt, it will move faster, making more of it to melt. After enough water is in the ditch, open the end and let it fall on over the edge of a cliff or mountain into the hydro-electric plant. If no mountain is handy, it will be necessary to construct a man-made water-fall. This water will turn the turbines and create electricity to run the heating coil.

<u>Examples Score 4</u>

Dig a great drainage ditch to catch the run-off. Then, set up a great water network all over the glacier which would spray water over the freshly fallen snow and melt it. This would, also, melt through the glacier in many places, cutting it up into smaller pieces. This would go on until the glacier had completely melted and, of course, would stop its advance by lightening the amount of bulk ice in it.

Build an artificial lake of immense size, miles from the town and build a dam for it and a hydro-electric plant. Having estimated the amount of water in it, with immense laser beams, melt the front of the glacier carefully so as not to cause a flood. As the water reaches the dam it will build up slowly and eventually flow out, at which time the hydro-electric plant will convert the energy into electricity. This water could be run off by a series of canals. It could be used for irrigation, reservoirs and the remaining water would have to be distributed to areas further away. The electricity created would not only enable all this to be done electrically, but it would, also, eventually more than

pay for the amount spent to stop the glacier. The water in the glacier would be melted as far back as they wanted, and if kept melted back into a valley, they could melt it as it rises out of the valley.

Item 86. He is a professor of English at a large University.

Examples Score 0

Qualified - He would also know about science if he is well educated.

Not qualified - He would be more interested in spelling and grammar.

No effect - He could read about it in a book.

Examples Score 1

Qualified - This shows he could write a good article if he wanted to. Anyone could look this up in a book.

Not qualified - All of his training would be in literature and not much is written about science in literature.

No effect - Glaciers may be a hobby of his or he might not even know what one is.

Examples Score 2

Qualified -

Not qualified - Since he is a professor of English people wouldn't believe him even if he wrote a good article because he is not a scientist. Also, he would have to copy most of the material and wouldn't know what was correct and what was wrong.

No effect - Most English teachers read a lot and have special interests. Maybe he would be interested in glaciers from some studies he had in college.

Item 87. His books are often quoted in newspapers and magazines.

Examples Score 0

Qualified - he must be good.

Not qualified - A lot of people are quoted.

No effect - Maybe he is quoted by communist papers.

Examples Score 1

Qualified - If he is quoted about glaciers then he is qualified.

Not qualified - Magazines and newspapers don't usually quote scientists, so he must not be one.

No effect - This doesn't say what kind of newspapers and magazines.

Examples Score 2

Qualified -

Not qualified -

No effect - Maybe his books are about glaciers, then this would make him qualified if his theories are right, but if his books are about cowboys and Indians, then he would not be qualified. Since the question didn't say, I don't know what it is, so I'll say no effect.

Item 88. He spent a year at Thule Air Force Base in Greenland.

Examples Score 0

Qualified - He would have seen a lot of snow.

Not qualified - The Air Force isn't interested in glaciers.

No effect - Maybe he was there in the summer.

Examples Score 1

Qualified - He would have been able to observe glaciers first hand.

Not qualified - If he spent the year at the base he would not see the real glacier because of all the buildings.

No effect - Maybe he was from Florida and couldn't stand to get out in the cold, or maybe he looked at them a lot.

Examples Score 2

Qualified -

Not qualified -

No effect - This doesn't say why he was sent to Greenland. If he was sent to study glaciers it would help a lot, but if he was sent as a mechanic and wasn't interested in glaciers it would not help.

Item 89. This was the first article written by the author published in Scientific American.

Examples Score 0

Qualified - It has good articles.

Not qualified - He wouldn't have enough experience.

No effect - It might be good or bad.

Examples Score 1

Qualified - The Scientific American is a good magazine, and would check the article before publishing it.

Not qualified - If this was his first article, then he wouldn't have a lot of experience and wide recognition as an expert.

No effect - Maybe he has published a hundred articles somewhere else.

Examples Score 2

Qualified - The Scientific American prints easier versions of complicated theories and research, so this would mean he had a good reputation among scientists and was persuaded to print it where more people would see it. How many people read the International Glacierologist Monthly?

Not qualified.

No effect - The article should be judged by itself and not on how many things the author has published. He may have put a lot of research in it or just copied from a couple of books.

Item 90. His books are often quoted by other glaciologists.

Examples Score 0

Qualified - He is well known.

Not qualified - They probably don't believe him.

No effect - He could be president of the Glacier Society.

Examples Score 1

Qualified - This shows that other glaciologists know of his work and respect his opinions.

Not qualified - This means they don't believe him and are trying to show where he is wrong.

No effect - He may be studying the small mountain glaciers and not know much about the polar ice caps.

Examples Score 2

Qualified - This shows that his work is well known. If other

glaciologists quote him, then his beliefs are important and must be known if we are to understand glaciers even though some may be trying to disprove his theory, because the article says there is a broad range of estimates of the amoung of water in glaciers.

Not qualified -

No effect - This shows one of two things. Either he is an expert on glaciers or something that is closely related. If he is an expert about glaciers then he is qualified; if about something related, then not qualified.

Item 91. He won a Nobel prize for his work on the geology of Haiti.

Examples Score 0

Qualified - This is science.

Not qualified - He wouldn't know about glaciers.

No effect - This only means he was interested in rocks.

Examples Score 1

Qualified - This means that he is a highly skilled scientist and would write about a lot of areas in science.

Not qualified - There are no glaciers in Haiti, so he wouldn't have worked with them.

No effect - Geology is affected by glaciers, so if Haiti ever had a glacier he would have studied about them, but if it never had a glacier, then he wouldn't have.

Examples Score 2

Qualified - In becoming a recognized expert he would have had to learn about all the things that effect geology. Since glaciers affect the land, he would have studied about them and would be qualified.

Not qualified - This shows that he would have specialized in tropical lands and would not be working with glaciers. If he wrote about glaciers he would be talking about things he hadn't studied much and might make false statements.

No effect -

Item 92. He is a professor of meteorology (the study of weather) at a large university.

Examples Score 0

Qualified - He knows a lot about weather.

Not qualified - He wouldn't be an expert on glaciers.

No effect - Maybe all he knows about is clouds.

Examples Score 1

Qualified - Glaciers play an important part in weather, so he would have studied about them.

Not qualified - He would be more interested in how they effect temperature and wind and not about how they form.

No effect - Maybe he is a specialist that doesn't have much to do with glaciers, or maybe he specializes in the effect of glaciers on weather.

Examples Score 2

Qualified - Since the distribution of water in solid and liquid forms are most important in weather he would be interested in how the distributions are changing. Also, since mountain glaciers play a role in local weather he would have to know about them.

Not qualified - This time would be spent in non-glacier regions since there are no universities on glaciers so he would not be able to study them. He would just teach how to predict weather, not about glaciers.

No effect -

Item 93. His name is not included in the Directory of American Scholars.

Examples Score 0

Qualified - Scholars are philosophers.

Not qualified - If he is qualified he would be.

No effect - Probably a lot of other scientists' names aren't either.

Examples Score 1

Qualified - This is an organization of writers and poets so I think it shows he is probably a scientist.

Not qualified - He must not be very well known if this includes names of scientists, therefore, I would think him less qualified.

No effect - If he is young he wouldn't be well known enough yet, but could still write a good article

Examples Score 2

Qualified -

Not qualified -

No effect - Maybe he is not an American, or if he is maybe he is still in college. This doesn't say if he is working on glaciers or what.

Item 94. He is an amateur mountain climber.

Examples Score 0

Qualified - He could get to them.

Not qualified - He should be expert.

No effect - Maybe he likes to walk.

Examples Score 1

Qualified - He would climb through glaciers on the high mountains, and so would see them first hand.

Not qualified - If he is an amateur he would know the difference between snow and a glacier.

No effect - Doesn't say why; is it for the exercise or to study glaciers?

Examples Score 2

Qualified -

Not qualified - An amateur mountain climber would stay away from the hard areas like glaciers and would go just to be outside and to look at birds, flowers and rocks. Professional climbers go to lead scientists to study things and would learn a lot themselves.

No effect - This doesn't say what his main job is; if it was related to glaciers he would have to climb some mountains secondarily to his job, but if he has a job not related to glaciers, then it wouldn't help.

Item 95. He is a recognized expert in the field of electronics.

Examples Score 0

Qualified - He would know about a lot of things.

Not qualified - This has nothing to do with glaciers.

No effect - No relationship

Examples Score 1

Qualified - He could build instruments to measure glaciers.

Not qualified - All of his training would be in electronics and engineering, so he would know little about glaciers.

No effect - Maybe he studied glaciers as a hobby.

Examples Score 2

Qualified - In order to study glaciers you have to be able to build, or at least use, complicated instruments. Since most instruments used would not be used anywhere else, you would about have to be an expert in electronics to study glaciers. Also, they have to be able to repair their own equipment.

Not qualified - He would not be qualified because he would not be interested in how the weather is affected by glaciers. All he would be interested in are things like missles and computers.

No effect - Maybe he is recognized as an expert by his neighbors because of what he learned from glacier instruments when studying, or maybe he is known because of his work in the space program.

Lisbon Earthquake

Item 81. Suppose that you had held an influential position in Lisbon and that you were opposed to moving the capital to Rio de Janeiro. Other influential people supported the proposed move. Briefly outline the reasons you would have used in an attempt to convince them that the capital should remain in Lisbon.

Objective: The objective of this item is to force the students to project into the situation immediately following the earthquake. Since the article emphasizes the instability following the earthquake, the subject should point out stabilizing effects of leaving the capital in Lisbon or the further disruption which would be caused by moving the capital.

Examples Score 0

It shouldn't be moved just because of the earthquake.

I would not be against it because things were so bad in Lisbon.

Lisbon is a good location near the sea.

Examples Score 1

There must have been some good reason why the capital was first located in Lisbon, for example, location or climate, and these reasons would still apply even though the city was destroyed. The city could be rebuilt.

An earthquake or other disaster can happen anywhere, so moving would not insure that the new capital would not be destroyed. Also, there would be a lot of mix up in moving.

The capital should stay in the mother country unless some kind of take-over had taken place. Instead of moving the capital to another country moving it to a different city.

First, I would mention tradition. Then, I'd tell them it was idiotic to have the capital of a country in another country and it would no longer be the Portugese Empire but the Brazilian empire.

Examples Score 2

The main arguments I would use are:

1. All of Lisbon was not destroyed.
2. There were important documents and papers to be found and transferring everything to Rio de Janerio would cost more than rebuilding the capital.
3. This is where important officials live and they would have to move.

Building a new capital in Rio would take longer than in Lisbon, because material from damaged buildings would be available. Also, there would be no government for months while the officials were moving to Rio since ships didn't have radios then.

Lisbon belongs to the mother land and the capital should not be moved to a colony. It's closer too of major powers and commerce than Rio. It's closer to the center of trade in the world and more accessible to traders and merchants.

I would oppose the move because:

1. Poor transportaion and communication.
2. Portugal would become a colony and Brazil the 'mother country.'
3. South America is underdeveloped.

Examples Score 3

Lisbon had been the capital for so long, and it could be rebuilt. The economy of Lisbon would drop tremendously because you wouldn't have the trade and turnover of people you had had. Your particular business could be affected because the capital employs a lot of people. Your own personal life could also be affected in that many people would move and so many people have been killed that those surviving cannot stand to loose their friends and family as they need one another.

I. Lisbon remain the capital:
   A. Lisbon is the capital, known as the capital (worldwide), through all books and papers--Lisbon is known as the capital.
      1. Earthquakes strike all over, not only in Lisbon.
      2. If another city became the capital it would cause many political problems, would have to

move out their government and replace it with ours.

3. Rio de Janeiro--ground may not be as solid and good as Lisbon is now since the earthquake has hardened the foundation under Lisbon.
4. The move and starting over would cost more than staying in Lisbon.

Reasons for keeping the capital in Lisbon:

1. Rio de Janeiro is too far from the mother country.
2. We must show strength and determination by not giving in to a natural disaster.
3. The people of other countries and of our own country will think more highly of our determination if we stay in Lisbon.
4. The people of Lisbon will feel abandoned if we leave.
5. All Portuguese must pull together to completely rebuild our devastated capital.

<u>Examples Score 4</u>

The capital should not be moved, I feel, because to rebuild the city will take much work and many people. To move the capital would remove the many people which will be needed. It would also kill the spirit of the people who remained. There would also be a problem of transportation. To rebuild the city people need to be brought in, not taken away. The spirit needs to be raised, not lowered. If the city remains as the capital there will be a want and a need to rebuild it.

I. Capital should be in Portugal:
   A. Brazil is too far away.
      1. Lack of communications.
      2. Lack of unity.
   B. Capital of Portuguese people should be in their own country.
   C. Moving capital would weaken Portugal.

II. Brazil is unsuitable:
   A. Backward, vast and poor communications, Indians, Negroes, etc.
   B. Independence Movement in Latin America.

As a patriotic citizen of the great city of Lisbon it would be unthinkable to move the capital to foreign city thousands of miles to a country recently discovered and still full of hostile Indians. One must think of the cost

of such a rash move. One might also check into the background of the people who proposed such a move; it might prove interesting to see how much some of them would benefit by such a move.

Pombal has managed to reorganize our city. In time we may be better off than we used to be. Besides, the capital has always been there. It is in a better location here. It will be an awful lot of trouble to change and besides, what if we decide to move it back again. "Mark my words", Lisbon will soon be rebuilt into a greater city than ever before. Pombal is a born organizer. Besides, we loose control if it is moved to Rio de Janeiro. We will no longer be an important city and many of you may lose your wealth and power.

Lisbon Earthquake

Item 82. Suppose a political cartoonist wished to make the philosophy of Leibniz appear foolish. Draw or describe a cartoon which would accomplish the above purpose.

Objective: The purpose of this item is to force the student to translate a set of abstract relations, given in the reading passage, into a concrete presentation.

Examples Score 0

A figure (man or animal) labeled Leibniz. Examples: The devil, a beggar or tramp, mule, a deformed man, a dog, etc.

A picture of a person eating the earth with the caption "The World is suited to our desires and appetities."

Examples Score 1

Drawing or description of Lisbon in a state of destruction with a caption quoted directly from the passage. Examples: "Your particular misfortune is nothing, it contributes to the universal good" and "It will be repaired by natural means."

Drawing or description showing a person in distress with the caption, "What is, is right." or "It is God's will that you should die."

A drawing or description of a person living in luxury while others are living in poverty with captions, "What is, is right," "This is the best of all possible worlds", or "I can't help the poor because there is as much harmony as possible."

Examples Score 2

A drawing or description of a person committing a crime with the caption, "I can't help myself, it is determined I should be a criminal." "Don't be upset by my crime, it contributes to the general good." "I am helping to keep balance by taking from the rich."

A drawing of Lisbon in a state of destruction, with Liebniz saying, "It will be repaired by natural means." with a second drawing dated several years later showing ruin covered with weeds and trees.

A drawing or description of a person in distress and another person looking on saying, "If you die it will be for the universal good." "If it is in the master plan, you will be okay."

A drawing or description of a man about to be executed with the caption "You are upsetting the natural balance by removing me from the world," "You are meddling with the plan of the world."

Examples Score 3

A drawing or description of contrasting action that followers of Leibniz and Voltaire would take following a disaster, and showing that Voltaire's followers would make the most progress.

A drawing or description of modern day leaders discussing international affairs with the caption "We believe Leibniz philosophy, therefore we can do nothing to stop the inevitable destruction of the world through atomic warfare because the plan has been set." (Any modern day problem could be substituted.)

A drawing or description of a civilized city contrasted with an uncivilized village indicating that civilization followed Voltaire and uncivilized followed Leibniz.

Examples Score 4

A drawing or description showing prosperous individuals accepting Leibniz with the caption "Leibniz is good." with poor individuals saying "Voltaire is good." or other indications that Leibniz philosophy cannot be accepted by the repressed.

A drawing or description of political leaders discussing the building of a hydro-electric dam with some proposing rejection because if it were in the plan for the world, a natural dam would exist, or other indication that man should not use modern engineering to change the major topographical aspects as an area.

A prosperous nation which follows Leibniz denies food to an underdeveloped country, and the prosperous nation has a depression because of lack of trade while another prosperous nation following Voltaire gives food to a similar nation and becomes more prosperous because of trade with the formerly depressed nation. (This could be shown by pictures of people talking or by dock scenes with ships sitting idle for the "Leibniz" nation and with the ships moving for the Voltaire nation. Then later on you would show "bread lines" for the "Leibniz" nations and "work lines" for the "Voltaire nation."

Item 83. The article implies that the nobility opposed Pombal's reforms. Suppose you were a member of the Portuguese nobility. In a few sentences describe why you would have opposed Pombal.

Objective: The subject must indicate that the reforms described in the article were reforms which would weaken the power structure in existence. Since descriptions of medieval life are assumed to be common knowledge among high school students the task is one of noting the consequence of reform to the power and prestige of the nobility.

Examples Score 0

I would not oppose his reforms myself if he was making the country a democratic one rather than a dictatorship.

Pombal's reforms were too modern for the time that he lived.

Examples Score 1

I would want to keep my position in life. I would not want to share my riches with others. I like the old ways; why should I want to change?

Pombal was for a fast (radical) reform which was not the answer for such a disaster area. A more conservative leader should have been appointed.

It would have taken too long to build it up. They were all ready well off and didn't care about the other people.

Pombal would be trying to redo something in a short time which took years to build. It would probably destroy the faith the people had in the clergy and nobility if someone came in and told them that the way things were wrong and began to change them.

Examples Score 2

During this period a form of feudal system was still in effect. Changes Pombal would induce would destroy my position as a member of the nobility and thus losing my special privileges. Doubtlessly, I would refer to retain these privileges, and would oppose Pombal's reforms.

He had too much power. The people would become educated and revolt. He changed the taxation and civil law for his own good. If the reforms are made the rich (nobility) will have to pay for them.

He is controlling everything his way, at least with us the control was in more hands. He's taking the rights away from all the nobility and is not making religion an important part of one's life.

Examples Score 3

As a member of the nobility I feel that Pombal is out of his mind to think he can do this. After all we are better than these lowly peasants who roam the streets. Why should we do anything for them. They should be doing for us. His reforms will never come through because they are beyond our reach.

Pombal wanted things to happen quickly. "Good things take time," I say. He wanted a modern city in an old country. The people werehappy before, they liked their way of life. I feel we must get 'modern' gradually, after all moderniza-tion is not an overnight thing.

Examples Score 4

Pombal is becoming too powerful; he is over-stepping his ground. We no longer hold any power. All he wants to do is modernize. We should be content with the way things are instead of always trying to modernize. Many of his methods won't work; he's just wasting his time and ours. One thing for sure, he's got too much power, and some of it used to belong to us.

Pombal's reforms are ruthless and a product of dictatorship. We do not wish to see the lower classes rise up to equal us. We are doing well without modernization. He is not doing things by the consent of the people, so why do it..if they do not agree. Maybe they do not want to break tradition or custom.

I, as a member of the nobility, think that Pombal's reforms are a sin in the eyes of God and a curse to this country. The Bible says, "Ye shall be content with what you have and ye shall not covenant." This, Pombal is doing, only under the name of a different type. It is like the wolf in sheep's clothing.

Pombal is an arrogant opportunist who caught our government off-balance in the crisis of an earthquake. He is destroying the culture and great traditions. We have sought for a long time. His dictatorial powers will soon envelope all that we have known and our once great capital city will be no more than a factory run by a whip-cracker.

Item 84. Events are described differently by different people according to the views which they hold. Describe how a Lisbon priest might have described the earthquake and its immediate aftermath.

Objective: This item is designed to test the students creative writing ability. The basic ingredients given in the reading passage must be reconstructed into a vivid description of what it might have been like to have witnessed the destruction. The projec-tion into the views of a priest implies that the description would be one of compassion rather than one of detached observer.

Examples Score 0

He would have told about the destruction.

He would have described it about like everyone else. It was a disaster. He would also try to help in getting the town back on it's feet.

It was a terrible thing.

This is God's will, and we must not lose faith.

Examples Score 1

A priest might say that God makes everything happen for a good reason. The people must not have been in accord with Divine Law or the city would not have been destroyed. There was so much evil in the city that they brought the destruction on themselves.

God has punished us for our sins. The terrible jolts of the earth and the mass ruins are a warning to all mankind. Be good to the Lord and he will be good to you.

The earthquake was very bad and killed many thousands of people, most of whom were in church; but, I believe that it was God's will that it be done, and I am not afraid of such things.

It is a force sent by God to show people that they have sinned. They are homeless and have no food because they have not believed in Him.

Examples Score 2

He might have started by saying that the disaster was the will of God and that God was punishing the people for some wrong they had done. He may have described the disaster as some of the disasters mentioned in Bible times, such as destructions which God brought upon various cities then. The peoples' death was their punishment, according to the priest.

He would have told how the people in the churches reacted, scared, running around, etc. He would have told about going out into the ruins and looking for survivors and

helping people that were hurt and wandering around. He would have told of setting up relief posts for the needy.

The poor people of my parrish have suffered so much with the quakes and fires and all. It is terrible to think of the deaths and wounds the poor people of Lisbon have suffered. I pray that this shall never happen again on this earth. There is chaos and looting everywhere; law and order must be established and emergency shelters set up to care for the sick.

Examples Score 3

I was celebrating All Saint's Day in my church when around ten o'clock a slight tremor shook the church. About two minutes later another quake was felt, and it seemed as all hell broke loose. Suddenly I realized the roof was going to cave in; and I yelled to the people to get out as quickly as possible as I shoved some children out by way of the vestry. Just as I got out a third quake hit and the church crumbled as if it was made of clay. For hours and hours I worked with the others trying to find survivors, giving last rites and praying for the dead when necessary.

The earth shook and made terrible noises as if the end of the world had come, then the buildings collapsed and the candles in the houses which were lit in reverence, the very candles lit in praise, burned our people and their homes. In the end there was a terrible silence, the only sound was the moaning of the maimed and the stench of the dead. All I could do was to ask mercy from the Lord for my people.

The first tremor I hardly felt. The second tremor, being considerably worse, shook the surrounding foundations. The walls and floor opened into giant cracks. I noted that the roof appeared solid as yet. During the next few minutes the room was a scene of chaos. When the second tremor ceased, I indicated that the room be cleared. However, before anyone could move the third shock hit. I remember seeing cracks in the ceiling, but nothing further. When I came to, I found myself some distance from what had been a church.

Examples Score 4

I remember the morning; it was very nice, with birds singing and the sun shining. Mass would be right on schedule at 9:00. People were coming in and trying to find a seat next

to the cool stone walls of the church. My alter boy went out and lit the candles as I came out I saw people standing up. They were very ragged and had no shoes, but they all had smiling faces. Well! As the Mass was near the end there was a slight tremor. I remember because one of the candles fell near the alter; then, after ten or so minutes it happened as I stood next to the door near the alter, the whole stone structure seemed to collapse. The next thing I knew someone was helping me away from the rest and I blanked out. When I came to and started to help with the digging it was a terrible sight to look at.

"The day of reckoning has come," "The sky is falling," and m^ny other things were shouted that day. But, I knew none of these things were true, for the Lord said he would come on a cloud, and that none would know the time. He said that all eyes would be on him, and yet I saw nothing. I called upon the members of the church and calmed them; we all went into prayer. The church started to crumble and then everything went black. I know not what happened after that. I just pray the Lord be with those who left the world that day.

Item 85. The Lisbon article is approximately three pages in length. Attempt to summarize the entire article in five sentences.

Objective: The major objective of this item is to force the subject to condense the article, leaving out all but the most significant parts. It tests the student's ability to view the article as a congruent whole and put from the whole the major aspects of its structure. This summary must be more than a mere copying of a sentence from the reading passage since this will result in a complete summary.

Examples Score 0

A disasterous earthquake started a philosophical war.

There were three quakes, a capital was destroyed.

The entire article deals with an argument on whether the earthquake was for the good of the people or for the bad.

Examples Score 1

On All Saint's Day, in a place called Lisbon, an eathquake struck three times. Many lives were taken and many were

hurt, so there came some aid from Great Britain and Spain, and also Portugal. During this confusion, the capital was to be moved and a man was appointed Minister of Foreign Affairs and War by a king called King Jose I. The man's name was Pombal; this man had many new and different ideas, and he planned to rebuild this place. During this, when all these happenings about this place was taking place, their spending different ideas about this happening, many different important men had different ideas about this matter, and they all discussed this among themselves.

The Lisbon earthquake occurred on November 1, 1755, killing many people and destroying towns 20 leagues away. After the earthquake a man named Pombal started to rebuild Lisbon into a more modern society. There were many philosophers who made many statements about the earthquake, how it was a good thing, how it was brought about through bad government and so on.

The article tells about the earthquake itself. The needs of the city and the nations that helped. The rebuilding of Lisbon by Pombal. Then, the affect the earthquake had on the philosophers of that time. The article reminds us that nature's calamities are still with us and still threaten us.

Many people sinners, good, bad, and medium, were all praying on Saints Day. Most of them probably sinners. 2. Earthquake shakes the town three times, kills thousands of people. 3. Survivors are shocked and anew government is brought in; much conversation about the cause and the ending of earthquake. 4. People rebuild city, new modern way and organized new governments. 5. Everyone lives happily after that.

Examples Score 2

In 1755 a giant earthquake of 3 tremors came on All Saints' Day in Lisbon, Portugal and killed many people. Fire, a tidal wave, looting and falling stone helped. It began a philosophical discussion which went back and forth between Voltaire, Kant vs Liebniz, Rosseau. The discussion concerned: free will, nature, predestination and mankinds actions.

On November 1, 1755, All Saints Day, an earthquake destroyed a leading city, killing thousands of men, women and children. 2) As a result of this disaster, food, clothing, money and other aid was sent from many allies. 3) A new leader rose in power, rebuilding the city into a new and better one. 4) The disaster caused men like Voltaire and Rousseau

to comment and express their ideas about the earthquake and about life in general.

Examples Score 3

In 1755 a great earthquake hit the city of Lisbon, demolishing the city, killing about 30,000 people, and leaving the surrounding area in economic turmoil. Pombal became dictator and immediately began reforms to rebuild the city. Pompal was successful in rebuilding, stopping pillage and bringing about his reforms. After the earthquake, many noted philosophers took note of the incident and prepared statements concerning their beliefs. No matter how progressive civilization becomes, natural and man made calamities will occur creating varied philosophical ideas.

On All Saints' Day in the year 1775 a violent earthquake occurred in Lisbon, Portugal at 10 o'clock in the morning while most of the population was attending church services. Fires, resulting partially from the great number of All Saints Day candles, earthquake shocks, and tidal waves destroyed all but the stronger new part of Lisbon. Pombal a politically strong man, became a dictator of a country without a capital. Aid was given by Britain and other nations while Pombal tried to prevent plague and rebuild Lisbon into a modern city. His reforms were opposed by the nobility and clergy and the whole topic of the Lisbon earthquake became a subject of great controversy among the great philosophers of the time, with Rousseau and Voltaire arguing from a more progressive point of view, and with Leibniz taking the stand of "this is the best of all possible worlds" predestination type of philosophy.

Examples Score 4

On November 1, 1755, an earthquake occurred in Lisbon that completely destroyed the town and almost all of the people in it. From this confusion and despair rose a dictator, Pombal, whose social reforms and plans set Lisbon back on its feet and set Portugal on the road to modernization. Meanwhile, European philosophers and writers were discussing the earthquake and why it happened as a matter of world Philosophy. Leibniz and Rousseau lead the belief that it happened because that was how it was meant to be and nothing could change, while Voltaire argued that man is responsible for his own destiny and the destiny of his fellow man. There is no clear cut proof of which

is right, but the discussion still goes on today and disasters are still occurring.

The Lisbon Earthquake was one of the worst that has ever occurred. Thirty thousand persons were killed in this disaster which lasted about twenty minutes. Pombal, a dominant figure in Portuguese politics, who was appointed Minister of Foreign Affairs and War, made himself a leader of that country by helping the people of Lisbon find food and shelter, and then rebuilding the city according to a master plan. Because of the catastrophe, great philosophers such as Kant, Leibniz, Voltaire and Rousseau, began to theorize what was or was not good for man and his world.

Item 86. He wrote a book condemning Leibniz.

Examples Score 0

Qualified - He would know about the earthquake.

Not qualified - He would not write about him.

No effect - He might agree with Voltaire.

Examples Score 1

Qualified - To write the book he would have found out about the earthquake and the damage it did.

Not qualified - He would try to make Leibniz look bad and would not give other viewpoints.

No effect - Leibniz's philosophy is only a small part of the article and he could know about the other parts.

Examples Score 2

Qualified - In such a book he would have to compare him with other philosophers of his time. Since most of the philosophy of the time was about the earthquake, he would have learned a lot about it.

Not qualified - Much of the article was about the earthquake itself and should be written like history. He would not know about the destruction of the earthquake from reading

about what people thought about the meaning of the disaster.

No effect - Maybe in studying about Leibniz he read about the earthquake and got real interested in it and wrote about the earthquake, but he might have not been interested. Also, maybe he was condemning Leibniz's philosophy when applied to modern times so it would not have any relation to the earthquake.

Item 87. He is a professor of world history at a large University.

Examples Score 0

Qualified - He would know about it.

Not qualified - Because this is in European history.

No effect - Some professors do not write articles.

Examples Score 1

Qualified - Because this is a major happening in history because of all the reforms so he would know about it.

Not qualified -

No effect - World History is a big subject and he might not be interested in Portugal.

Examples Score 2

Qualified - Because he would have a lot of knowledge about how it changed history and would have books to look up what he didn't remember. The philosophers in the article are studied in history so he would know about them too.

Not qualified.

No effect.

Item 88. He is a retired junior high school librarian.

Examples Score 0

Qualified - He would have a college degree.

Not qualified - Librarians do not study history.

No effect - This article could have been written by anyone.

Examples Score 1

Qualified - Most of this article is taken from books and he would know how to look them up.

Not qualified - Junior high school libraries do not have many books on this subject, so he would not even know about it.

No effect - This shows that he has the ability to write the article, but doesn't mean he was interested in it.

Examples Score 2

Qualified - Librarians spend most of the time working with books so he would have a lot of information already. After working with books all of his life he would know a lot about how to write and could easily look up the information and write an article.

Not qualified - To write an article like this you would have to know a lot about all the philosophers who wrote about the earthquake and about the earthquake itself. A librarian would just take what someone else had written about these things and shorten them.

No effect - This says nothing about his interest and knowledge of the history and philosophy of the 17 and 18 hundreds. If he wanted to he could write the article if he went to a big library to look up the information, but it doesn't say he did.

Item 89. He has never been outside the United States.

Examples Score 0

Qualified - Most authors haven't, but write good books.

Not qualified - He would have to see it to write about it.

No effect - People who write about Mars haven't been there either.

Examples Score 1

Qualified - If he went there he would hear about the earthquake through accounts that had been handed down and would not write a true story.

Not qualified - He would have to go to Portugal to find old records.

No effect - It happened so long ago that going there wouldn't be much help.

Examples Score 2

Qualified -

Not qualified -

No effect - Since the earthquake happened so long ago it would not do any good for him to go. All of the ruins have been removed, rebuilt or grown over with grass, so he could stay here and find out just as much.

Item 90. He lived in Portugal until he was 25 years old.

Examples Score 0

Qualified - He could have seen it.

Not qualified - He wouldn't know the philosophy.

No effect - He might not have heard about it.

Examples Score 1

Qualified - He would have studied about it in history.

Not qualified - He would be more interested in the damage than the philosophy which followed it.

No effect - He might have studied about it in history, but this wouldn't be much help.

Examples Score 2

Qualified - The earthquake would be studied more in Portugal than in any other country and there would be a lot of old records for him to study.

Not qualified - Because if he had heard of it, he would only hear about the death of his ancestors and would be biased against Leibniz who said their death was good.

No effect - This would have no effect because even though he might have studied it in history he would be no better qualified than I am to write about the Chicago fire which was in our history book.

Item 91. He wrote this article while teaching social studies in a small town.

Examples Score 0

Qualified - This is part of social studies.

Not qualified - Because he wouldn't be smart enough.

No effect - He could have lived in a large city most of his life.

Examples Score 1

Qualified - Because the earthquakes would be of interest to him and he would have time.

Not qualified - This would take reading and a small town wouldn't have a big library.

No effect - If he was interested he could send off for books, but if he wasn't then he wouldn't.

Examples Score 2

Qualified - Teachers are always looking for things for students to do for outside work and being a social studies teacher he might have found some information on it. After he found the information, he would have been able to write a good article because of all the writing he had to do in college.

Not qualified -

No effect - He may have taken a lot of courses in European history in college and found a lot of stuff to write and took the job so he would have time or he might have just been smart enough to get a degree and couldn't get a job in a big city which have good schools.

Item 92. He is a well known philosopher and has written many books and articles.

Examples Score 0

Qualified - This has a lot of philosophy.

Not qualified - He probably never heardof Lisbon.

No effect - His books may be about science.

Examples Score 1

Qualified - He would have a good background from his studies about Voltaire and Leibniz.

Not qualified - If he is well known he probably is writing popular philosophy and is not interested in old philosophies.

No effect - This doesn't say if he is also interested in history and natural disasters of the past.

Examples Score 2

Qualified - In all of his work he would have studied a lot about the philosophers in this article and used their views.

When figuring out their philosophy he would have noticed they were talking about the earthquake and would have learned about the earthquake.

Not qualified - He would be writing against one of the philosophers in the article so would change what happened after the earthquake to fit his view.

No effect - He is a good writer but so are a lot of other people; so the only thing is if he liked to write about the earthquake and this doesn't say if he did or didn't.

Item 93. He is a well known author of historical fiction.

Examples Score 0

Qualified - He knows how to write.

Not qualified - Fiction isn't true.

No effect - He could still write the truth.

Examples Score 1

Qualified - Historical fiction is based on true happenings, so he would know about the real happenings.

Not qualified - He would make up things to make a better story.

No effect - He might be writing about the Old West or Europe.

Examples Score 2

Qualified - This shows two things, one he knows how to write interesting stories and two, he knows a lot about history to make his stories show how life was then.

Not qualified - He would not be qualified because all he would be interested in is what happened to people since this would make a good book. He would not know about the philosophy because this would not make good fiction.

No effect -

Item 94. Pombal and he were good friends.

Examples Score 0

Qualified - He would have seen it.

Not qualified - He would be against reform.

No effect - Pombal had a lot of friends.

Examples Score 1

Qualified - He would have lived during this period and could have gone to see the ruins if he didn't live through the earthquake.

Not qualified - He would be biased in favor of Pombal's reforms and would not have written the truth about them.

No effect - He might not even know how to read and write.

Examples Score 2

Qualified - He could have seen the damage, talked to survivors and gotten his opinions about Pombal's reforms directly from him. He could have also talked with the philosophers if he wanted to.

Not qualified - If Pombal and he were good friends, he could not have lived long enough to see the outcomes of all the philosophy which lasted 100 years.

No effect - Maybe he was a personal friend or maybe he was his friend and was also in the government. In either case it doesn't say if he was interested in philosophy which is what all of the writing was about.

Item 95. He has been blind since birth.

Examples Score 0

Qualified - He could better understand suffering.

Not qualified - He couldn't see the destruction.

No effect - He could listen to records.

Examples Score 1

Qualified -

Not qualified - He would have a hard time learning to read Braile well enough to find all the facts.

No effect - He could use Braile.

Examples Score 2

Qualified -

Not qualified -

No effect - All of the remains would be gone now anyway so that wouldn't do him any harm. With all of the stuff in Braile and on records he could do the research as well as anyone if he could get the council of the blind to pay for it.

Stages of Economic Growth

Item 81. "Suppose that you are a national leader in Australia. Describe some of the things that could be done to enable the country to reach maturity."

Objective: The objective of this item is to test the students realization that even though Australia is in high mass-consumption, some phases of its economy are weak. The reading passage implies that it reached high mass consumption by neglecting to develop its industrial capabilities, therefore, things could be done to make a sounder economy.

Examples Score 0

I would increase social security payments.

I would start a program to export raw materials.

Leave things alone because Australia is in High Mass Consumption and will reach Maturity in a few years.

I would go out and conquer more land so we would have more raw materials.

Examples Score 1

Start factories of all kinds. Build new schools, and start a new and improved education system. I would try to educate everyone, and teach them skills besides farming.

Have more technical colleges and schools. New factories built, Agreements with other countries on importing and exporting. Develop complex tools and chemicals more transportation as railroads and waterways.

Australia's industries could be expanded and trade stepped up. New industries could be made to help her more self-sufficient. Technical education could be emphasized more. Australia's resources could be developed and made available to other countries.

(1) Various housing projects could be built to compensate for urban growth. (2) Increase educational institutions in all aspects. (3) Encourage industrial investment and

also further development (4) Strive for better methods and means of international trade.

Income could be reinvested in the industries therefore strengthening technological maturity. (Emphasis should be placed on the production of consumer goods in order to fulfill demand for them). Expansion of factories provides more skilled jobs and decreases the unskilled labor.

Examples Score 2

I would conduct a campaign to make the people realize how important investment is in making our nation a strong and influential nation. I would also start schools to educate people about the new machines and their usefulness and also how to work them. I would also support any sane movement for organization of labor.

The first thing that would be needed to be done, would be to increase the technical skills of the people. Schools would have to be started that taught how to make, use, and repair complex machines. The people would be encouraged to start factories. Also, the rate of exporting and trading with the rest of the world would have to be stepped up. Better working conditions and higher salaries should be instituted to make the factories desirable to work in.

Australia, in order to reach maturity, will have to develop in many directions. It must strive to enter international trade. It must improve its technology. It must produce more of its own goods and tax foreign goods to promote home industry. It must mechanize farming to allow more work in industry. It must be increasingly aware of the fact that industrialization is all-important.

Examples Score 3

I. I would increase the variety of goods produced. II. Maybe step backward a bit and pay more attention to improving technology than to society, decrease benefits put to industrial aid. III. Give aid to new businesses. IV. Give aid to schools to train technologists. V. Disconcern myself with world power.

(NOTE: point no. II and V indicate recognition of the fact that Australia is in High Mass Consumption at the present time.)

As leader, I would invest more money in scientists and their experiments. The money could be used most in experiments dealing with new chemicals that could lead the country toward producing many of the things which are necessarily imported. Also, more money would go to persons experimenting with new industry. I would also want to invest much of the money to industries that will bring in more money for the government. Education would play a very big part, also. I would have better schools built, more technological schools than anything else.

First, I would start improving technology through schools and research, probably with government help. In order to decrease the number of people needed in agriculture and therefore increase the number of potential skilled industrial workers I would promote research for better methods of food production. I would also encourage the national investment to increase through better banking service and increase real wages of the workers.

Examples Score 4

If I were a national leader in Australia, I would try to get the government to support scientific research for technical improvements in farming and manufacturing methods. New industries of the desired kind, such as those manufacturing machine tools, chemicals or precision instruments, could be given tax exemptions. Tariffs could be raised to protect the infant industries. There would be social legislation to protect the workers; minimum wage laws and social security. Treaties with other nations and the formation of reciprocal trade agreements would help us to secure a place in international trade.

Since it has gone through take off and passed through high mass consumption, an inbetween stage needs to be organized. It would first be beneficial to mechanize agriculture both to provide food for home use and possibly export. Mechanization would also free agricultural workers to work in newly established industries. I would make protect these new industries with import taxes. Another extremely important factor would be increased and improved educational facilities, providing both general and technical skills for the population.

Item 82. Political campaigns are based on issues that are important to the citizens. If you were running for a national office in a country which was in maturity what promises would you make in an effort to win the election?

Objective: The objective is for the student to realize that a country in maturity is moving toward the goals of high mass consumption. Therefore, the emphasis would be on the availability of consumer goods and include a program to combat depression and insure future economic security for individuals.

Examples Score 0

New industries would be established to make your life easier with their new conveniences. The men will be treated fairly with the management of their companies. Our needs will be less expensive with most of our goods being made in our country.

(1) A world leading power. (2) New technological advance. (3) Higher middle class and middle class almost equal to the present wealthy class. (4) Make our country better.

Build new factories with better machines to employ more people.

Examples Score 1

A society in maturity has a working force mainly in industry. Therefore, promises which would benefit manufacturing and industry would be in order. Plans favoring unions, higher wages, and increased security would be inclined to capture the majority of the votes.

I would promise the factory workers higher wages and greater social security benefits. I would reduce the cost of living.

I would promote the school system. I would make promises that would help the country socially and politically. Further state welfare. Make possible better jobs and more jobs for the people.

I would increase social legislation designed to redistribute income, promise shorter working hours, and increase social security. I would also promise to make consumer goods more available.

Examples Score 2

I would promise a buildup of the nation's defenses and offensive ability; I would promise social legislation to better working conditions; provide some sort of a welfare program, whether it be social security, medicare, or some such thing. I would also try to increase the productivity of my nation by increasing trade relations with other nations, and encourage more education to train skilled people to put out more products.

I would promise the people greater fruits of the labor they are performing. Higher real wages, better standard of living, and increased social security. I would appeal to the poor with promises of urban renewal and increase employment. To the factory workers I would give a higher minimum wage and better working conditions. To the every increasing middle-class I would offer improved educational facilities and more leisure time.

First of all I wouldn't make any promises -- I would propose ideas. 1. reduce unemployment. 2. Clear slum areas. 3. Aid for older people. 4. More money for education.

I would promise more old age benefits and a social securities program expansion. I would promise to strengthen the countries economic social and political powers through more frequent and closer ties with foreign countries. I would promise to do everything possible to increase the leisure time of the average man in accordance with more pleasant and less working hours to parallel with the advancement of machinery in industry.

Examples Score 3

1. Set up a graduated income tax system, social security, and unemployment insurance. 2. Create a favorable attitude toward our government by sending technicians to other countries and at the same time try to influence their governments. 3. Education for all through high school and give college scholarships for bright students.

I would promise to consider all of the people and not just a chosen few. I would introduce a bill that would increase the minimum wage and improve the working hours. I would also stress the need to have strict health standards in public buildings and restaurants, also health regulations

in rented apartments. Thus, improving the living conditions of the laboring class. There needs to be a special scholarship fund set up so students who could not afford to stay in school could do so. I would investigate our Social Security system, and bring about changes so that it would increase the benefits.

Since the people wouldn't be interested in increased technology, I would concentrate more on social benefits. I would set up a Department of Cultural Affairs to help strategically create cultural centers around the U.S. I would have the Department of Interior start a War on Ugliness. Probably more directly important to the people would be an effort by the government to increase the security and welfare of the majority of the people by federal retirement funds, unemployment insurance, slum clearance, and housing.

Examples Score 4

Issues important to citizens in a pre-high mass consumption stage would relate job and general economic security. Therefore I would promise a more adequate pay scale, minimum wage law, minimum government standards on working conditions and safety facilities. For economic security I would set up an unemployment insurance plan for those who, through no fault of their own, cannot find work, and a retirement insurance plan for all citizens. Government restrictions on big business providing for retraining of workers who are or might be displaced by industrialization. Increased federal personal loans for education and housing. I would also set up a system to insure savings accounts to encourage use of banking facilities and free more money for investment.

If I were running for office I would promise to improve and increase education and educational facilities. The setting up of junior college would be good for those students who could not afford to leave home to attend a large university. Being in the maturity stage--it would be of interest to the people how I felt about unemployment insurance and retirement insurance--both of which I would promise to support. I would also support labor unions which would be set up to improve the workers security and working conditions. I would introduce a bill to set up a plan for slum clearance and urban renewal, thus helping to establish better living conditions for the working class. Another step in this direction would be low rent federal housing projects. Another point I would bring out would be a tax law to help redistribute income - - in order to help the "little man."

Item 88. "The list of manufactured goods of countries in take-off and in high mass consumption are similar in many ways and are different in many ways. Make two short lists of manufactured goods, one representative of the manufactured goods of a country in take-off and one representative of the manufactured goods of a country in high mass-consumption. How are your two lists alike and how are they different?"

Objective: Both economics have the basic extracting and processing industries, and manufacture heavy equipment and necessities and refine food products. However, only high mass consumption has the benefits of technical advancement in industrial production and the consumer goods and luxury products.

Examples Score 0

"cotton
steel
manufactured raw materials

cotton shirts
shovels
manufactured raw materials put into use."

"take off
steel goods
iron goods
cloth

High mass-consumption
steel goods
iron goods
cloth

In the take-off manufactured goods are just beginning their production. The amount is smaller than in high mass consumption. In high mass consumption there would be a greater amount of goods produced."

"Take-off
coal
iron
steel
aluminum
raw materials
plastic

cloth (imported)
cars
clocks and watches
jewelry

They are alike in that they are trade with other countries, nations, and states. They are different in that one (high mass) is consumer goods and one is a new export sector.

Examples Score 1

Take-off
coal, iron, books, textiles, steel products
High-mass-consumption
machines, glass, syn. materials home utilities

The two lists should be basically the same, except the latter deals with more highly developed items. The take-off stage develops just the basically needed items, while in the mass consumption, there is a demand for many various items.

Take-Off	High Mass-Consumption
Tractors	Television sets
Plows	Automobiles
Horse Carts	Lawn Mowers
Stoves	Electric Ranges

My lists are alike in the function of the goods but the main difference is in this technological advances found in the High Mass-Consumption list. The horse cart in the Take Off has been motorized in High Mass-Consumption. The stove in Take Off has been given electricity in High Mass Consumption. The main differences are in the advances that have been made.

Take-Off	High Mass-Consumption
Farm Machinery	Clothes
Farm Products (food)	Factory Products
Factory Goods (cars, machinery)	Machinery
	Cars
	Metals
	Defense Materials
	Transportation Materials
	Consumer Goods (household furniture)
	Educational Needs

In take-off they're starting on farm goods and factory goods, but by the time you (country) reach maturity and high mass-consumption you start on bigger things like defense, transportation, household goods, educational needs, etc.

Examples Score 2

Take-Off	High Mass-Consumption
Coal	Electrical Equipment
Iron	Televisions
Tools	Agricultural Products
Chemicals	Greater Transportation Means
Agricultural Goods	

In a way the two lists are alike in that they include things that the country uses for advancement, but High Mass-Consumption is concerned more with consumer goods than those of any worth to government or industries.

Take-Off	High Mass-Consumption
Food	Heavy Machinery
Processed Natural Resources	Processed Natural Resources
(Some manufactured goods such as various farm implements, ship for exporting things, etc.)	Luxury Items
	Foods (some)

The lists are similar in the nature of the articles although their quantity and quality greatly vary. Also, the luxury items represent a vast amount of varied industries which the take-off period undoubtedly lacked.

Take-Off	Mass-Consumption
Canned Goods	Luxury Items
Textiles	More Varied Transportation Vehicles
Transportation Vehicle	Textiles
Marine Goods	Dry Goods
	Marine Goods

In Mass Consumption there are more and more varied goods than in Take-Off. Also, Mass Consumption tends to keep the same things in as there were in Take-Off, but to find different and better ways of making them.

Examples Score 3

Take-Off	High Mass-Consumption
Steel	Medicine
Machinery	Caviar
Tools	Sports Cars

Goods manufactured during the take-off stage are those which enable increased industries and new methods. They are building blocks. During the high mass-consumption, the economic system has reached a balance and there is no longer a need for increase of industries on a large scale. Manufacturers produce more luxury items to be consumed by the public instead.

Take-Off	Mass Consumption
Coal	Steel
Steel	Clothing
Farm Machinery	Transportation
Clothing	Synthetics (rubber)
Transportation (cars)	

Both lists are alike in many ways because the needs of the two periods are similar in many ways such as the need for steel to build factories, cars, machinery and the need for clothing and food products for the growing population. They are different in the amount produced and the type produced e.g. synthetics such as rubber and likes for use in the home, for transportation and for defense and the amount required for the population will grow with the needs of the population and the wealth of the population. The country in the Stage of high mass-consumption will need the more complex synthetics and the larger amount of clothing, etc., for its population and will be able to supply it.

Take-Off	High Mass-Consumption
Corn, Wheat, Rice, Cereals	Foods (processed)
Farm Equipment	Petroleum Goods
Crude Oil	Byproducts of all Types
	Machinery
	Luxuries

The lists are alike in the manner that the goods are basically the same except they have been developed to a far greater degree in the high mass production stage. Technology and development of resources has greatly increased.

Take-Off	High Mass-Consumption
Clothing	Clothing
Food	Food
Household Goods	Household Goods
Farm Machinery	Farm Machinery
Raw Materials (iron,coal)	Raw Materials
Industrial Machines	Machines
	Cars
	T.V.
	Household Appliances
	Colored Princess Phones
	Toys
	Luxury Items

These lists are similar in that they both contain industrial and raw materials. These goods are necessary in any society but especially in take-off where they form the basis for the economical change. But in Mass Consumption state these goods

are necessary to maintain this achievement but also you have a lot of consumer goods which were not present in take-off. These come from an increased income and money flow and a demand for luxuries and conveniences.

Item 84. According to the chart, not all countries spent the same amount of time in each stage, nor did they progress through the stages in the same order. Briefly describe conditions under which a country might be able to skip or combine stages.

Objective: The student should realize the importance of external forces in economic development. Forces which radically alter typical development include colonization, foreign aid in wars, discovery of a natural resource important in the international market, etc.

Examples Score 0

If a country is in one of the lower stages and they have a political revolution then they could skip a stage and really start everything rolling. They can also combine stages maybe if they start out fast in one stage before they complete it they have already started in the next stage.

If the country was very rich in natural resources, good conditions, favorable public support and strong capable leaders.

If they have a lot of income then it would be possible for a country to combine such stages as take-off and high mass-consumption. If they have a smart president or someone to run the government well they could skip stages.

If a country starts progressing at a very fast pace, progression might not cease and new and better things will be discovered. The maturity stage could possibly be skipped by entering the last stage because of not needing the time and progression covered in the maturity stage.

Examples Score 1

If they were helped by foreign countries and had sufficient resources.

"(1) War
(2) Post War
(3) Industrial Revolution
(4) Take over by another country."

This depends upon the condition of the rest of the world, and the conditions of the country itself. For example--the country might be involved in a civil war, or there might be a world war in progress. Also, the inventions of another country might be used to help the country skip a step.

Due to Canada's close ties with Britain the first take-off I suspect that Canada was able to progress from a specifically wheat producing country to one of industrial status by:

(1) Being close to Britain
(2) Exporting her wheat to other countries such as U.S. and Britain for materials and leadership in her progression.

<u>Examples Score 2</u>

A country which was not technically advanced might have a product which was internationally important. So that while increasing its manufacturing it could also develop exports greatly. Maturity would be reached while take-off was going on because the country would have basic and technical industries to advance at the same time. A country which received aid would be able to more quickly move through the stages of growth. A new central government which was democratic rather than a dictatorship would seek more equality among the people and better national improvements.

Countries that are developed under the guidance and aid of highly developed nations, tried to skip and combine stages. These newly-developing countries have the benefits of years of progress and knowledge at their disposal. They can receive aid from their parent country. With such benefits, these countries may easily develop quickly by skipping or combining stages.

A new unexpected technological advancement, foreign aid in the form of technical advisory committees, finances, etc., a rapid change over in the form of government are all effects that might enable a country to progress rapidly by skipping a stage of development.

Examples Score 3

Certain countries may be able to skip or combine stages because of interior economics and/or external factors. The invention of a special tool or the discovery of a natural resource that was needed by another country and that country would help with the refining of the resource. You could also build up your foreign trade this way. If another country were to give you financial help or technical help--such as Peace Corps. War could be an external factor that would help a country to skip or combine stages.

There are several possible reasons that a country might skip or combine stages--briefly these are:

(1) An underdeveloped country might receive economic and technical aid from another country.
(2) A migration may occur from a more developed country for political or other reasons.
(3) A needed natural-resource may be found and developed by an outside interest.

Australia might be a country that could skip the traditional period for these reasons:

(1) England could start with trained workers in Australia.
(2) Money not needed to build military could be used for industry, education, etc.
(3) Could benefit from natural resources discoveries.
(4) Foreign aid from other countries (Example of combining)

Examples Score 4

There is always a possibility that a war, especially with a more advanced society, could change the order of the stages. Also, discovery of a new chemical or product that is in demand by countries involved in the space race such as a rocket fuel, or a positive cure for cancer.

Influence of another more advanced society by colonization, technical aid, or financial assistance such as is the case with Great Britain and Canada, and the U.S. Peace Corps.

In war and other stress conditions a country might move by necessity for survial to a higher economic stage without going through intermediate stages.

Or a new discovery--such as a brilliant technological advancement or a new source of revenue such as petroleum which is desperately needed by other countries might enable a country to advance skipping usual steps.

A country which is a possession of a more advanced society which has been colonized by the mother country would provide both a ready market for the country's goods and a source of technical aid to enable rapid development.

In summary, elements which initiate a skipping or merging of stages usually comes from external factors: war, colonization, and natural resources.

Item 85. Suppose that you are president of a country which is poor, underdeveloped, and in the transitional stage. A wealthy nation has offered aid in any one of the following forms: machines, professional and technical personnel, surplus food, or household goods and clothing. Choose the one form of aid which you think would produce the greatest long term benefits for your country, and tell why that form would be more beneficial than the others.

Objective: The student should recognize the impact of basic necessities for a developing nation and the relative value of each, for present relief as opposed to future security. The purpose of the item is to force the student to choose between several desirable and necessary forms of aid and to justify his choice over other forms of aid.

Examples Score 0

Machines, because this could help us back on our feet so we could help ourselves.

I would take surplus food, because I wouldn't want my country to do without food.

Machines because there everybody could get a job and move up.

The second: Then the poor people who haven't had any food in a long time could rest while they worked.

Examples Score 1

I would say professional and technical personnel because if you have skilled, intelligent people working in your country, the country would grow and become stronger. It would greatly develop probably in all fields. Then you wouldn't rely on another country for aid.

Surplus food. Without food how could a person run a machine, learn from high-trained personnel. A person has to have food before he can do anything.

Professional and technical personnel could teach the people the correct procedures for production. These newly taught people could teach other people and so on. The trained people from the other country could immediately begin improving the underdeveloped country They could help further advance the new ideas of that country.

Machines because it can help to increase the rate of production while the others can be obtained after they have reached for a good start.

Examples Score 2

I would take the professional and technical personnel. They could show the country what steps to take to get self-supporting and industrialized. The other goods would help at the moment, but would not produce any long-term effects. The machinery would certainly help, but not if no one knew how to run the machines. Therefore, the technological help would be needed.

The machines would be the most beneficial. The professional and technical personnel would benefit the country for only a short time, but when the people reached the training capacity of these people, they would be at a standstill with no way for advancement. The surplus food would eventually run out and the country would be where it started from. With the proper machines, the people would be able to furnish this food continuously. The same with household goods or clothing. These could be done with proper machines. These machines would soon advance into more complex machines and the country would still benefit.

Professional and technical help would bring the long term benefits machines would bring short term results, but by using the technical help we could learn to build our own machines.

Professional and technical personnel would be my choice because they could tell us how to get what we want to get. Also, if we got machines then they would only hold us for a short while. I think that we were informed of all these things then we would be better off in the future and we could do things for ourselves.

Examples Score 3

I would choose the machines. With the machines industries could be begun, trade would be increased, products would be increased. There would be a greater flow of money throughout the country and it would be fairly evenly distributed so that the little man would benefit as well as the bigger ones. A more or less chain reaction would be begun. The people would work with the machines producing more products, which would produce more trade, which would yield more money, which would give more to the rich as well as the poor, thus making more food for all, more conveniences, and making the educational aspect greater; education would make professional and technical personnel. Thus, through machines you can have all the other things that were offered.

Professional and technical personnel--I'd want my country to stand on its own as soon as possible and the greatest long range aid available is education. My people would learn skills and be able to pass them onto their countrymen-once having attained this technical knowledge we would no longer need outsiders--once taught, we could teach others. Our learning of technical skills, vocations, and professions would not break down or wear out--it would be the greatest productive investment we could make in our greatest resource--our people.

(I'm an idealist) (sic)

I would choose professional and technical personnel. This would benefit my country the most because it would teach them how to grow for themselves. Machines would be of little value without the technical no how and surplus food would be of immediate value but would soon be used up and we would return to our former state. Household goods and clothing would help but it would do nothing to increase our knowledge, it could only make us better off with what we already know. Personnel would be the wisest choice because with it we will be able to prepare for the future while increasing present standards by ourselves.

Examples Score 4

I believe that professional and technical personnel would be most beneficial. Machines by themselves would be useless, surplus food would make a healthier nation but not more prosperous in other ways, and while household goods and clothing would not give the people the initiative to work to gain things for themselves, professional and technical personnel, on the other hand, would teach the people to use their own capabilities which would in the long run actually benefit the people and their families and future families.

The benefits of sending in skilled personnel would be long range, extending to future. Each generation could build on the preceding.

I believe I would take the professional and technical personnel. They would be able to teach my country new ways of doing things. I would be able to start schools with excellent teachers, and teach my people the ways of the new industrial society.

I would not accept the machine, because there would be no one qualified to teach the people how to run them. If I take the professional and technical personnel, they could teach the people. Since in most countries which are in the traditional stage the people are mainly farmers, surplus food would be no good. The people would grow their own. The same goes for the household goods and clothing. It would help the people, but it would not advance the nation.

Evaluation Sample Responses

Item 86. In the mid-1930's Japan was better prepared for war than was China.

Examples Score 0

More trust. The chart says so.

Less trust. Since neither country had reached maturity they should have been equally well prepared.

No effect. China was better prepared because of her extremely large population.

Examples Score 1

Greater trust. Japan had almost reached maturity while China wasn't in Take-Off, so Japan would be more industrialized.

Less trust.

No effect. Rostow doesn't see war as an important aspect of economic growth, therefore, Rostow's theory would say nothing about preparing for war.

Examples Score 2

Greater trust. According to Rostow's chart, Japan was in drive to maturity. This would mean she was developing manufacturing and could build some materials for a modern war. China was traditional stage and would have very little industry, therefore, would not be able to supply the goods for a modern army.

Less trust.

No effect. The chart shows that Japan had almost reached maturity and, therefore, according to his theory, had more industrialization than China which was not even in Take-Off. But from the above statement one can't infer that Japan necessarily had more industrialization since it could have been receiving aid from Germany or started to train its army earlier. Therefore, this statement wouldn't necessarily affect my trust.

Item 87. Rostow is an economic advisor to the President of the United States.

Examples Score 0

Greater trust. If the President trusts him, he must be good.

Less trust. I think the President has the wrong economic theory.

No effect. Presidents can make mistakes.

Examples Score 1

Greater trust. Presidents usually pick only the most influential men in their fields as advisors.

Less trust. As advisor he would be most concerned with U.S. Economics and would not have time to develop a theory for all countries.

No effect. As advisor he might be working in such a narrow field that he could not see the "big picture."

Examples Score 2

Greater trust. The Presidents' advisors are notusually full time staff members so he would not have to be working in a special field and Presidents usually ask for advice from several outstanding experts in a given field.

Less trust.

No effect. Advisors are usually called for specific problems such as possible effects of a new farm support program. This man may be a good agricultural economist and not be a general theorist or he may have been called because of his knowledge of the effects of granting foreign aid to underdeveloped countries.

Item 88. Several great civilizations have developed and disappeared in the course of world history.

Examples Score 0

Greater trust. This shows that countries constantly change.

Less trust. He said countries always get better.

No effect. Rostow doesn't say that countries always continue once they are formed.

Examples Score 1

Greater trust. Since technology was not developed they could not get to transition so they were more susceptible to take over.

Less trust. Rostow's theory implies that once a central government has assumed control over the economy (become civilized) it will continue to expand.

No effect. His theory could not be applied since they did not have technology.

Examples Score 2

Greater trust. Rostow states that development depends on use of technology to permit the economy to expand. Since technology was not available, these countries could not expand economically. These civilizations were very much tied to agriculture and conquest so if they had a crop failure or were attacked by a strong nation they would be weakened and fall.

Less trust. If you assume that history repeats itself then the great civilizations of today will also fall through internal deterioration, therefore, Rostow's theory is inadequate because it does not include a stage for this.

No effect. Most of the great civilizations did not fall because of economic weaknesses. Since only traditional stages were reached we cannot assume Rostow to be wrong. From history we know that they fell because of outsiders invading them.

Item 89. A particular state in the United States has 80 per cent of its population engaged in agriculture; however, the wealth and state governmental control are in the hands of several dozen large land owners.

Examples Score 0

Greater trust. This state is in the traditional stage.

Less trust. The United States is in high mass-consumption.

No effect. Such a state does not exist.

Examples Score 1

Greater trust. That state is in what Rostow would call the traditional stage, but we can assume that this state benefits from the higher stages of the other states.

Less trust. Rostow said that the wealth of countries in high mass-consumption is fairly well distributed and few people are farmers and this is not true of this state.

No effect. Rostow was concerned with the classification of entire countries and did not say that all sections had to be classified the same.

Examples Score 2

Greater trust.

Less trust. According to Rostow's theory this state would be in the traditional stage. In order for the economy to expand workers must be free to work in non-agricultural jobs. With 80 per cent of the population in agriculture, there would not be enough people to work as school teachers, policemen, mechanics, barbers, social workers, and all other jobs in high mass-consumption. Therefore, regions of the country would have to be classified differently and Rostow classifies entire nations.

No effect. This state could be the leading agricultural state of the nation and if it used modern methods many times the population of the state would be fed by its output. Also, since farm machinery is expensive the landowners would have a lot of wealth in terms of net worth but would have most of their wealth invested in machinery.

Item 90. There are several poorly developed countries that have been completely controlled by strong dictators for several decades.

Examples Score 0

Greater trust. They must be traditional countries.

Less trust. In that amount of time they should have developed.

No effect. Since it is a traditional society controlled by a dictator, Rostow's theories doesn't apply.

Examples Score 1

Greater trust. Rostow's conditions for the traditional society are fulfilled since the presence of a dictator doesn't mean that the other more important conditions don't exist.

Less trust. Countries in the traditional stage are, according to Rostow, supposed to have regional governments. Since they have dictators, I would have less trust.

No effect. Rostow's theory doesn't say how long a country will remain in any one stage.

Examples Score 2

Greater trust. Since "poorly developed" can refer to a range of development and since a dictator can be considered a 'strong central government', these countries can be assumed to be on either preconditions for take-off or late traditional. In either case Rostow's theory is supported.

Less trust. The very fact that dictators can inhibit the growth of these countries shakes Rostow's theory since he assumes that all countries grow. Also, the presence of a strong dictator in a basically traditional society goes against one of his major conditions, regional power.

No effect. Rostow's theory is an economic theory not a political theory. It doesn't deal with the conditions which stop the growth of a country such as power-hungry dictators but only with the conditions which lead to growth.

Item 91. Mr. Rostow is a college student who wrote the article as an assignment for a course in business economics.

Examples Score 0

Greater trust. If he is a college student he is pretty smart.

Less trust. College students have a lot of theories.

No effect. He might be an expert and still be in college.

Examples Score 1

Greater trust. If he were in college, he could go to the library and look up a lot of different theories, then write his own.

Less trust. As a college student he wouldn't have the time to spend on developing a really good theory because this would take years.

No effect. He might be a graduate student who has studied economic theory for years or he may be a freshman and this is his first course in economics.

Examples Score 2

Greater trust. He could use the library to gather lots of facts, then he could talk to his professors and get their ideas, then he could write his theory and check it against several countries by looking them up to see their production.

Less trust. I would believe his theory less because he would just write anything to get a good grade. He probably read a book and copied most of it changing some parts so it wouldn't look like he copied. Also, if this were just for one course he probably isn't really an economic theorist and may not even be majoring in economics.

No effect. The main thing to consider is how true is the theory, because if it is true then it wouldn't make any difference who wrote it. Just on the surface it would look like he is not qualified but he may be a retired government official who knows a lot of economics taking courses just for fun.

Item 92. Several new nations are being created in Africa.

Examples Score 0

Greater trust. Because this shows the world is getting more developed.

Less trust. Rostow says nothing about new nations.

No effect. Because they were not on the chart.

Examples Score 1

Greater trust. Because this shows that the nations are expanding beyond their own borders and creating new nations.

Less trust. Because new countries are not created according to Rostow they develop from loosely organized clans when technology demands it.

No effect. Because as the tribes become more civilized, they will form countries for better government. You would have to wait a 100 years to see how they progress to tell about the theory.

Examples Score 2

Greater trust. In high mass-consumption Rostow states that one thing a country does is "pursuit of external power." This would mean that they would gain power by developing new nations. This is necessary to have a place to export goods and to get raw materials.

Less trust. Rostow indicates that countries evolve out of loosely organized groups. If creating a new nation could overcome this natural evolution then his theory would be in error.

No effect. Because we cannot tell at what stage they are being created or by whom and for what purpose. Probably the new nations are being formed to help fight surrounding tribes so they would be more military than economic purposes.

Item 93. Mr. Rostow has two college degrees.

Examples Score 0

Greater trust. Because he has to be smart to get two degrees.

Less trust. Probably they are not in stages of economic growth.

No effect. A lot of people who write articles only have one degree.

Examples Score 1

Greater trust. I assume he studied economics or else he wouldn't be writing an article on it.

Less trust. They probably are not in economics so it makes no difference how many degrees he has.

No effect. What are the two degrees in? If in economics, he would be qualified, but what if they are in mathematics?

Examples Score 2

Greater trust. A person with two degrees has to have the ability to read and write good papers. They would have had the chance to use the library and study a lot of theory. One other thing with two degrees he would know enough not to write on something he didn't know anything about.

Less trust. From the way the article was written, it sounded like the person majored in History and if he did I would not trust his theory any more than anyone else. If he were an economics major the article would have talked more about money value exchange rates, gross national products, and real earning powers.

No effect. You do not say what the degrees are in so I will say they are not in economics. In this case he would have the ability but would not have the interest or motivation.

Item 94. Mr. Rostow wrote his theory and drew the chart in the reading passage in 1930.

Examples Score 0

Greater trust. Because he could see into the future.

Less trust. Times have changed since then.

No effect. He could write it anytime.

Examples Score 1

Greater trust. Since the chart extends 30 years after 1930, he must have had the theory worked out pretty good.

Less trust. He could not have predicted World War II in 1930 so its effects on economy would not be known and may have changed things.

No effect. Because his theory was general of all countries he could have written it anytime. The chart would just show general trends.

Examples Score 2

Greater trust. Since the best of my knowledge the chart is right then he would have had to have things worked out pretty well to predict in 1930. Since this is just a short summary I would think he uses a lot more to predict future economy than what is given in the article.

Less trust. Every country except U.S. and Canada have changed since 1930. He would just have been guessing about all of these changes. He had to make up a time for Canada to reach maturity since it had not reached it but was in high mass-consumption. How could he know about World War II and how it affected the countries?

No effect.

Item 95. The United States has a surplus of agricultural products.

Examples Score 0

Greater trust. The U.S. should have surplus since its agriculture is efficiently run.

Less trust. A surplus doesn't indicate the "balance" mentioned in the article for high mass consumption.

No effect. Rostow doesn't mention anything about surplus.

Examples Score 1

Greater trust. The technology that produces a surplus in America is a characteristic of high mass-consumption.

Less trust. Because overproduction of resources doesn't fit the description of high mass-consumption which is supposed to have achieved a balanced production.

No effect. The surplus would be used to aid poor countries and would be deliberately planned.

Examples Score 2

Greater trust. According to the chart the U.S. is in high mass-consumption; since it has passed maturity one would expect that the farming technology would be advanced enough to produce a surplus to be used in trade and in national emergencies.

Less trust. If the country had really achieved a true balance as indicated by the redistribution of income then we would not have a surplus because it would be given to the needy people.

No effect. Rostow's theory indicates that their major concern is providing for the welfare of all individuals and since this implies adequate food the presence of a surplus would not say anything about the theory unless it lasted a long period of time.

Reliability Estimates for Synthesis and Evaluation

Estimates of interjudge and alternate forms reliability were established for Synthesis and Evaluation for all forms. Reliability estimates were obtained for each Synthesis item; however, only total score reliabilities were estimated for Evaluation. Samples of answer sheets were drawn from one school. Answer sheets on which more than one Synthesis or two Evaluation items were omitted were not used. The remaining answer sheets contain a full range of responses but collectively they do not constitute a random sample.

These procedures were used to estimate interjudge reliability of the Synthesis items. Scores for each student were independently assigned and recorded by each rater. Correlations were computed for each pair of raters as an estimate of agreement between them. Thus for N judges, N(N-1)/2 correlations were calculated. These correlations were converted to the equivalent Fisher's z values, the mean z was computed, and it was converted to a correlation coefficient. The Synthesis item reliabilities appear below.

Synthesis Item Reliabilities

Item No.	Atomic Structure	Glaciers	Lisbon Earthquake	Stages of Economic Growth
81	.64	.56	.59	.68
82	.65	.59	.52	.54
83	.73	.54	.61	.67
84	.76	.70	.70	.64
85	.76	.73	.68	.68

Reliabilities for the Atomic Structure and Glaciers are based on three judges and 50 students. Reliabilities for the other two forms are based on six judges and fifty students.

To estimate total subtest reliability, total scores for Synthesis items for each student by each rater were tabulated and N(N-1)/2 correlations were calculated. Mean correlations were then calculated using the z transformation technique. This process was also used with the Evaluation scores. Total Synthesis and Evaluation interjudge reliability estimates appear below.

Interjudge Reliability of Synthesis
and Evaluation Subtests

	Synthesis	Evaluation
Atomic Structure	.89	.83
Glaciers	.72	.75
Lisbon	.71	.81
Economic Growth	.79	.72